

121/487

Assignment Report of WO 2006/078318

Thursday, July 29, 2004 6:49 PM

PCT/US2005/027239

Page 12

using Sanger sequencing method with Weighted residue weight table.

7150	ACCACATTACCACTTCTACTAGTACCCACCGTTACCATTTGATTTGAT	Majority		
7151	ACCACATTACCACTTCTACTAGTACCCACCGTTACCATTTGATTTGAT	cohl_a12.seq		
7151	ACCACATTACCACTTCTACTAGTACCCACCGTTACCATTTGATTTGAT	a909_a12.seq		
7160	7170	7180	7190	7200
<u>TGAGGTATCTTCTAATTTGATATCTCCTACTGGAAATAATGACAGGTTTA</u> Majority				
7210	7220	7230	7240	7250
7201	GGAGCTATCTTCTAATTTGATATCTCCTACTGGAAATAATGACAGGTTTA	cohl_a12.seq		
7201	GGAGCTATCTTCTAATTTGATATCTCCTACTGGAAATAATGACAGGTTTA	a909_a12.seq		
7260	7270	7280	7290	7300
<u>TGGTGATATTTTATAGCATCTGCTAAATGGGGTCAATATCAATGGAA</u> Majority				
7251	TGGTGATATTTTATAGCATCTGCTAAATGGGGTCAATATCAATGGAA	cohl_a12.seq		
7251	TGGTGATATTTTATAGCATCTGCTAAATGGGGTCAATATCAATGGAA	a909_a12.seq		
7310	7320	7330	7340	7350
7301	TCATATGGGTATAAATTTACCATTTGTAACCAACGCCACGGAAACGATA	cohl_a12.seq		
7301	TCATATGGGTATAAATTTACCATTTGTAACCAACGCCACGGAAACGATA	a909_a12.seq		
7360	7370	7380	7390	7400
<u>GCCATCTGGCATTGTCGGACGTCTCAGTAAGGCTGAATGGGAGCCATCGT</u> Majority				
7351	GCCATCTGGCATTGTCGGACGTCTCAGTAAGGCTGAATGGGAGCCATCGT	cohl_a12.seq		
7351	GCCATCTGGCATTGTCGGACGTCTCAGTAAGGCTGAATGGGAGCCATCGT	a909_a12.seq		
7410	7420	7430	7440	7450
7401	CATAGGAATCAGGGTCAGTAGGCTCATTACTTGTAAATTGCTGACCA	cohl_a12.seq		
7401	CATAGGAATCAGGGTCAGTAGGCTCATTACTTGTAAATTGCTGACCA	a909_a12.seq		
7460	7470	7480	7490	7500
7451	GAAGCATCCAATGCTGGCTTTCCATCTGTACCAACAGCATCATGCTGTA	cohl_a12.seq		
7451	GAAGCATCCAATGCTGGCTTTCCATCTGTACCAACAGCATCATGCTGTA	a909_a12.seq		
7510	7520	7530	7540	7550
7501	TATAATATGATAATCTCCAGCCTTCCCCAAATAGCTCTTAAATTGATA	cohl_a12.seq		
7501	TATAATATGATAATCTCCAGCCTTCCCCAAATAGCTCTTAAATTGATA	a909_a12.seq		
7560	7570	7580	7590	7600
<u>CTTGAGTTACAGCACCTGAAAAGTTATAAGGTCTAATACTACCATCTGG</u> Majority				
7551	CTTGAGTTACAGCACCTGAAAAGTTATAAGGTCTAATACTACCATCTGG	cohl_a12.seq		
7551	CTTGAGTTACAGCACCTGAAAAGTTATAAGGTCTAATACTACCATCTGG	a909_a12.seq		
7610	7620	7630	7640	7650
7601	TTAACATAATACCAACCGACTAATTGTAAAGCGTCTTTTACGTACTTGT	cohl_a12.seq		
7601	TTAACATAATACCAACCGACTAATTGTAAAGCGTCTTTTACGTACTTGT	a909_a12.seq		
7660	7670	7680	7690	7700
7651	CTTAGTTGTTGATCAACATTGAGAGACTAGTATCTGTCGTATAATAGG	cohl_a12.seq		
7651	CTTAGTTGTTGATCAACATTGAGAGACTAGTATCTGTCGTATAATAGG	a909_a12.seq		
7710	7720	7730	7740	7750
7701	CATCTTGTGAGTCGGGATCTTTATCTCGTGAATCATACCTATAATAA	cohl_a12.seq		
7701	CATCTTGTGAGTCGGGATCTTTATCTCGTGAATCATACCTATAATAA	a909_a12.seq		
7760	7770	7780	7790	7800
<u>TATGTACCTGAAGGATCTGGATATAATCCCTTGTAAATATCTGTATAATC</u> Majority				
7751	TATGTACCTGAAGGATCTGGATATAATCCCTTGTAAATATCTGTATAATC	cohl_a12.seq		
7751	TATGTACCTGAAGGATCTGGATATAATCCCTTGTAAATATCTGTATAATC	a909_a12.seq		

FIGURE 21K

122/487

Alignment Report of WO 2006/078318, n method with Weighted residue weight table.
Thursday, July 29, 2004 6:49 PM

PCT/US2005/027239

Page 13

CGG AATAC GAT CAC CATA ATG CAG T C TAA ATAGGTATCATCTGTTTG Majority
 7810 7820 7830 7840 7850

7801 CGG AATAC GAT CAC CATA ATG CAA ATCTAA ATAGGTATCATCTGTTTG coh1_a12.seq
 7801 CGG AATAC GAT CAC CATA ATG CAA ATCTAA ATAGGTATCATCTGTTTG a909_a12.seq

ATA ATT GGG CCT CCC GTT GGATCAATATTGACACGATATGTTACCTTTGC Majority
 7860 7870 7880 7890 7900

7851 ATA ATT GGG CCT CCC GTT GGATCAATATTGACACGATATGTTACCTTTGC coh1_a12.seq
 7851 ATA ATT GGG CCT CCC GTT GGATCAATATTGACACGATATGTTACCTTTGC a909_a12.seq

CAAC CTG CAT AGACT TTAAACATCATGAGGAGGCATAGTCTGTAAAGTC Majority
 7910 7920 7930 7940 7950

7901 CAAC CTG CAT AGACT TTAAACATCATGAGGAGGCATAGTCTGTAAAGTC coh1_a12.seq
 7901 CAAC CTG CAT AGACT TTAAACATCATGAGGAGGCATAGTCTGTAAAGTC a909_a12.seq

AAATACTTGTGTTGTGCTTGGTCTTTATACCATTTACCATCCC AAACAT Majority
 7960 7970 7980 7990 8000

7951 AAATACTTGTGTTGTGCTTGGTCTTTATACCATTTACCATCCC AAACAT coh1_a12.seq
 7951 AAATACTTGTGTTGTGCTTGGTCTTTATACCATTTACCATCCC AAACAT a909_a12.seq

ACC CTGGT CGACTAGGTTAGGTTGAACCGTTGTCTGATC GGGGGGCATAA Majority
 8010 8020 8030 8040 8050

8001 ACC CTGGT CGACTAGGTTAGGTTGAACCGTTGTCTGATC GGGGGGCATAA coh1_a12.seq
 8001 ACC CTGGT CGACTAGGTTAGGTTGAACCGTTGTCTGATC GGGGGGCATAA a909_a12.seq

GAGGACAAATTTGCTCATATAGAACATCCTTACTGGAAAATAGGAA Majority
 8060 8070 8080 8090 8100

8051 GAGGACAAATTTGCTCATATAGAACATCCTTACTGGAAAATAGGAA coh1_a12.seq
 8051 GAGGACAAATTTGCTCATATAGAACATCCTTACTGGAAAATAGGAA a909_a12.seq

CTCTGTATTATCAAGCGGATCTAKATATTTAACTCTTGTATGAATTACGT Majority
 8110 8120 8130 8140 8150

8101 CTCTGTATTATCAAGCGGATCTAAATATTTAACTCTTGTATGAATTACGT coh1_a12.seq
 8101 CTCTGTATTATCAAGCGGATCTAAATATTTAACTCTTGTATGAATTACGT a909_a12.seq

CATACCATACCACTAAAGTTCAAATACTTTGGTAGTCTCCATATTTA Majority
 8160 8170 8180 8190 8200

8151 CATACCATACCACTAAAGTTCAAATACTTTGGTAGTCTCCATATTTA coh1_a12.seq
 8151 CATACCATACCACTAAAGTTCAAATACTTTGGTAGTCTCCATATTTA a909_a12.seq

TCTGTAGTATTCTGCCATTGGCACTTTGGTATGCACTCGTTGTCT Majority
 8210 8220 8230 8240 8250

8201 TCTGTAGTATTCTGCCATTGGCACTTTGGTATGCACTCGTTGTCT coh1_a12.seq
 8201 TCTGTAGTATTCTGCCATTGGCACTTTGGTATGCACTAGTTGTAG a909_a12.seq

TGGGTCTGATCAAATAGTAATTATCTGGATATAAGCTTGTAGTATT Majority
 8260 8270 8280 8290 8300

8251 TGGGTCTGATCAAATAGTAATTATCTGGATATAAGCTTGTATATT coh1_a12.seq
 8251 TGGGTCTGATCAAATAGTAATTATCTGGATATAAGCTTGTATATT a909_a12.seq

TAACATTAATCTAGGTATTCTGTAAAGGTAAATCTGTCTGGTCTCCA Majority
 8310 8320 8330 8340 8350

8301 TAACATTAATCTAGGTATTCTGTAAAGGTAAATCTGTCTGGTCTCCA coh1_a12.seq
 8301 TAACATTAATCTAGGTATTCTGTAAAGGTAAATCTGTCTGGTCTCCA a909_a12.seq

GCACCTCCCCCTGTGCTGCTAAAGAGTATTGCCATCTAGTCTCTGTT Majority
 8360 8370 8380 8390 8400

8351 GCACCTCCCCGTATCTGCTAAAGAATAAGTGCCTATCTGCTCTT coh1_a12.seq
 8351 GCACCTCCCCGTATCTGCTAAAGAATAAGTGCCTATCTGCTCTT a909_a12.seq

GTAGAACGGATAATTGGATTCTCTTCCCTTTGGATAGAGTTTATT Majority
 8410 8420 8430 8440 8450

8401 GAAGAACGGATAATTGGATTCTCTTCCCTTTGGAAAGAGTTTGTAA coh1_a12.seq
 8401 ATAGAACGGATAATTGGATTCTCTTCCCTTTGGATAGAGTTTATT a909_a12.seq

FIGURE 21L

<u>C A T C T G A A T T T C T G T G A C T C C A T T G G G T A C T A T G A A C T C A C C C A A A Majority</u>					
	8460	8470	8480	8490	8500
8451	<u>C A T C T G A A T C A A C G G T A T C A T T A G G T A A A A T G A A C T C A C C C A A A</u> cohl_a12.seq				
8451	<u>C A T A A G G A T T T G C T T A A G T G A C T C C A C T G G G G A G T A T A A A C T C A C C C A A A</u> a909_a12.seq				
	<u>T A A C T C A T T C C T T A T G T T C C A G T T T G G T T A T T T C C A A C A T T G G T T A G G T A</u> Majority				
	8510	8520	8530	8540	8550
8495	<u>T A A C T C A T T C C T A A G T G T C C A A C T T G G T T A T T T C C A A C A T T G G T T A A A A T A</u> cohl_a12.seq				
8501	<u>T A A C T C A T T C C T G A T A T T C C A G T T G A A G T T A G G T A</u> a909_a12.seq				
	<u>A C G - C C A T G C A C C T G T C T T C C A T T G A T A G C C A T T C G G G G C T A A G G T T G T A C</u> Majority				
	8560	8570	8580	8590	8600
8545	<u>A C G - C C A T G C A C C T G T C T T C C A T T G A T A A C C A T T A G C G G G C G A A G G T T G T A C</u> cohl_a12.seq				
8536	<u>A C G - C C A C C G C A C C T G T C T T C C A T T G A T A G C C A T T C G C A G C T A A A G T T G T A C</u> a909_a12.seq				
	<u>C G T A T A G T C C T G T G T A G G T T C G G C A T C T G A T G C T C T A G T T C T A G G A A T C</u> Majority				
	8610	8620	8630	8640	8650
8595	<u>C G T A A A G T C C T G T G T A A G T A T C A G G C A T C A G A G C T C C A A T T A T A G G A A T A</u> cohl_a12.seq				
8586	<u>C A T A T A G T C C A G T A T A G G T T C G G C A T A T G A T G C A C T A G T C C A A G G A A A C</u> a909_a12.seq				
	<u>G T A G T A T T T G G T A A T G A A T C T C C G A G T A G C C C C T T T T T G C A A A T T T A T</u> Majority				
	8660	8670	8680	8690	8700
8644	<u>G T A A T A T T T G G C A A G G A A T A T C C C C C A T A T T T T G C A A A T T T A T</u> cohl_a12.seq				
8636	<u>G A A G A A G T T T G A T A A T A A A C C T C A G A G T A G C C C G T C T T T G C A A A T C T E A T</u> a909_a12.seq				
	<u>T G T G A T G A G T T T C T A T C A T A A T A A C A T T A A C G G A C A C T T G A A C C A T C G T</u> Majority				
	8710	8720	8730	8740	8750
8689	<u>T G T G A T G A G T T T C T A T C A T A A T A A C A T T A A C G A A C A C A C T T G A A C C A T C G T</u> cohl_a12.seq				
8686	<u>A G T G A T A A G T T T C T A T C A T A A T A A C A T T A A C G G A C A C T T G A A C C A T C G T</u> a909_a12.seq				
	<u>C T T T T A T C A T G A C A G A A G T T T C T G T C T C G T A T T A T T A A C T T T A A G C C A</u> Majority				
	8760	8770	8780	8790	8800
8739	<u>C T T T T A T C A T G A C A G A A G T T T C T G T C T C G T A T T A T T A A C T T T A A G C C A</u> cohl_a12.seq				
8736	<u>C T T T T A T C A T G A C A G A A G T T T C T G T C T C G T A T T A T T A A C T T T A A G C C A</u> a909_a12.seq				
	<u>G T C G G T A G T T T C A T T A A T A T C T T G T G T G T A G C G T C T G A T T A G A T A A</u> Majority				
	8810	8820	8830	8840	8850
8789	<u>G T C G G T A G T T T C A T T A A T A T C T T G T G T G T A G C G T C T G A T T A G A T A A</u> cohl_a12.seq				
8786	<u>G T C G G G C A A T T T T C A T T A A T A T C T T G T G T G T A G C G T C T G A T T A G A T A A</u> a909_a12.seq				
	<u>A G A T A G G G C T G A T C G T G T T A C T T G C C C T G C G T A C T C A T A T G T C T T T G C G</u> Majority				
	8860	8870	8880	8890	8900
8839	<u>A G A T A G G G C T G A T C G T G T T A C T T G C C C T G C G T A C T C A T A T G T C T T T G C G</u> cohl_a12.seq				
8836	<u>A G A T A G G G C T G A T C G T G T T A C T T G C C C T G C G T A C T C A T A T G T C T T T G C G</u> a909_a12.seq				
	<u>C A T C A G T A G G C A T T T I A T T A F C C G T T G C T G A T T G T T G C C A G T A G T T T A T C</u> Majority				
	8910	8920	8930	8940	8950
8889	<u>C A T C A G T A G G C A T T T I A T T A F C C G T T G C T G A T T G T T G C C C A G T A G T T T A T C</u> cohl_a12.seq				
8886	<u>C A T C A G T A G G C A T T T I A T T A F C C G T T G C T G A T T G T T G C C C A G T A G T T T A T C</u> a909_a12.seq				
	<u>G T G T A G G T T G T T G C G G G C A C C A A T G T G C A T A T A G C G T C G T A T C C T T</u> Majority				
	8960	8970	8980	8990	9000
8939	<u>G T G T A G G T T G T T G C G G G C A C C A A T G T G C A T A T A G C G T C G T A T C C T T</u> cohl_a12.seq				
8936	<u>G T G T A G G T T G T T G C G G G C A C C A A T G T G C A T A T A G C G T C G T A T C C T T</u> a909_a12.seq				
	<u>G G T C A A G A C T T G A T T A A A T C A A A G G C T G T C C C A C C A C T A G C G A G C T G T G T</u> Majority				
	9010	9020	9030	9040	9050
8989	<u>G G T C A A G A C T T G A T T A A A T C A A A G G C T G T C C C A C C A C T A G C G A G C T G T G T</u> cohl_a12.seq				
8986	<u>G G T C A A G A C T T G A T T A A A T C A A A G G C T G T C C C A C C A C T A G C G A G C T G T G T</u> a909_a12.seq				
	<u>A C C A C C C T G C A A A A G T A T A A C C T G G C C T T G T T G G A T C A T T A G G C T T A A T T</u> Majority				
	9060	9070	9080	9090	9100
9039	<u>A C C A C C C T G C A A A A A G T A T A A C C T G G C C T T G T T G G A T C A T T A G G C T T A A T T</u> cohl_a12.seq				
9036	<u>A C C A C C C T G C A A A A A G T A T A A C C T G G C C T T G T T G G A T C A T T A G G C T T A A T T</u> a909_a12.seq				

FIGURE 21M

	G T C G A A G C A G G G T T C G T G T T A A C C A C A G G A C G G A G G T G C A A T A T A G G T A A C	Majority
9110	9120	9130
9140	9150	
9089	G T C G A A G C A G G G T T G G C T G T T A A C C A C A C G A C G G A G G T G C A A T A T A G G T A A C	coh1_a12.seq
9086	G T C G A A G C A G G C T G G T C T G T T A A C C A C A C G A C G G A G G T G C A A T A T A G G T A A C	a909_a12.seq
	T C C T G T T G A T A A G T C G C C C T G T G T G A A T T C A A C A C C G T A C C G A T T C T T T A A	Majority
9160	9170	9180
9190	9200	
9139	T C C T G T T G A T A A G T C G C C C T G T G T G A A T T C A A C A C C G T A C C G A T T C T T T A A	coh1_a12.seq
9136	T C C T G T T G A T A A G T C G C C C T G T G T G A A T T C A A C A C C G T A C C G A T T C T T T A A	a909_a12.seq
	A G A C A G G A T A A A G A C T T A T G T C T G A A G A T A C A G G T A A T G C T T G A A T T T C T	Majority
9210	9220	9230
9240	9250	
9189	A G A C A G G A T A A A G A C T T A T G T C T G A A G A T A C A G G T A A T G C T T G A A T T T C T	coh1_a12.seq
9186	A G A C A G G A T A A A G A C T T A T G T C T G A A G A T A C A G G T A A T G C T T G A A T T T C T	a909_a12.seq
	G A T T C A G A A A G T G G G A G C T C C A T T T G A G T T T A G A C C A A C C T A C A A A T A A	Majority
9260	9270	9280
9290	9300	
9239	G A T T C A G A A A G T G G G A G C T C C A T T T G A G T T T A G A C C A A C C T A C A A A T A A	coh1_a12.seq
9236	G A T T C A G A A A G T G G G A G C T C C A T T T G A G T T T A G A C C A A C C T A C A A A T A A	a909_a12.seq
	T A A T G T T G A A G T G G G G T G G T G A A A C T T T A A A T G A A C T T A G A T C C C C G G T T C	Majority
9310	9320	9330
9340	9350	
9289	T A A T G T T G A A G T G G G G T G G T G A A A C T T T A A A T G A A C T T A A T C C C C G G T T C	coh1_a12.seq
9286	T A A T G T T G A A G T G G G G T G G T G A A A C T T T A A A T G A A C T T A G A T C C C C G G T T C	a909_a12.seq
	C G T T T C C T G A A C T A T C C A A T G G T A C T T G A T A A C G C T C C C A A A A T A C T T T A	Majority
9360	9370	9380
9390	9400	
9339	C G T T T C C T G A A C T A T C C A A T G G T A C T T G A T A A C G C T C C C A A A A T A C T T T A	coh1_a12.seq
9336	C G T T T C C T G A A C T A T C C A A T G G T A C T T G A T A A C G C T C C C A A A A T A C T T T A	a909_a12.seq
	C C A G C A G A A T C A T C A T A A A G T T A C T G T C G C C A C T T T A C T C A T A A A T G G	Majority
9410	9420	9430
9440	9450	
9389	C C A G C A G A A T C A T C A T A A A G T T A C T G T C G C C A C T T T A C T C A T A A A T G G	coh1_a12.seq
9386	C C A G C A G A A T C A T C A T A A A G T T A C T G T C G C C A C T T T A C T C A T A A A T G G	a909_a12.seq
	A C G A A C A T A A A T T T C T T T G T C A G T T A C A G G T A T T G G G T C A C C C A A A T	Majority
9460	9470	9480
9490	9500	
9439	A C G A A C A A A A T T T C T T T G T C A G T T A C A G G T A T T G G G T C A C C C A A A T	coh1_a12.seq
9436	A C G A A C A T A A A T T T C T T T G T C A G T T A C A G G T A T T G G G T C A C C C A A A T	a909_a12.seq
	T A A C A G G G T C A C C A T A C T T T C C A G T T A C T G A T A C A G G T A T A C C C A A C C A	Majority
9510	9520	9530
9540	9550	
9489	T A A C A G G G T C A C C A T A C T T T C C A G T T A C T G A T A C A G G T A T A C C C A A C C A	coh1_a12.seq
9486	T A A C A G G G T C A C C A T A C T T T C C A G T T A C T G A T A C A G G T A T A C C C A A C C A	a909_a12.seq
	T T A A A A T G C T C T C C T G C T T T A A T C G T C G G A A T C C C C A A C T T C T C C T A G A G C A	Majority
9560	9570	9580
9590	9600	
9539	T T A A A A T G C T C T C C T G C T T T A A T C G T C G G A A T C C C C A A C T T C T C C T A G A G C A	coh1_a12.seq
9536	T T A A A A T G C T C T C C T G C T T T A A T C G T C G G A A T C C C C A A C T T C T C C T A G A G C A	a909_a12.seq
	T T C T C C A T C T T T A A T T T G A T G A T G A A C T T G C A T A C C C T G A A G G C T G T C A	Majority
9610	9620	9630
9640	9650	
9589	T T C T C C A T C T T T A A T T T G A T G A T G A A C T T G C A T A C C C T G A A G G C T G T C A	coh1_a12.seq
9586	T T C T C C A T C T T T A A T T T G A T G A T G A A C T T G C A T A C C C T G A A G G C T G T C A	a909_a12.seq
	G G A A T T T A A A T C A G T T C C C G T C A T T A T T T G A A A A T G G T A A G T T A A C C T A	Majority
9660	9670	9680
9690	9700	
9639	G G A A T T T A A A T C A G T T C C C G T C A T T A T T T G A A A A T G G T A A G T T A A C C T A	coh1_a12.seq
9636	G G A A T T T A A A T C A G T T C C C G T C A T T A T T T G A A A A T G G T A A G T T A A C C T A	a909_a12.seq
	G G A A C T T C T G T A T T A T C C T G T G A A C A A T T G C A T A A A T G G G A G A T G A A T C	Majority
9710	9720	9730
9740	9750	
9689	G G A A C T T C T G T A T T A T C C T G T G A A C A A T T G C A T A A A T G G G A G A T G A A T C	coh1_a12.seq
9686	G G A A C T T C T G T A T T A T C C T G T G A A C A A T T G C A T A A A T G G G A G A T G A A T C	a909_a12.seq

FIGURE 2 IN

125/487

Alignment Report of WO 2006/078318, method with Weighted residue weight table.
Thursday, July 29, 2004 6:49 PM

PCT/US2005/027239,6

T G T T T A A A A G C G A C A T T A C I T G C T A G T G T T C T T A G T T T C T G C A G T A T C T T Majority				
9760	9770	9780	9790	9800
9739 T G T T T A A A A G C A A C A T C A C T G C T A G T G T T C T T A G T T T C T G C A G T A T C T T coh1_a12.seq				
9736 T G T T T A A A A G C A A C A T C A C T G C T A G T G T T C T T A G T T T C T G C A G T A T C T T a909_a12.seq				
T A G A T T T T A A T A C T T C T G T T G A C C A T C A T C T T T A A A G T G A A C A A C T T T A Majority				
9810 9820 9830 9840 9850				
9789 T A G A T T T T A A T A C T T C T G T T G A C C A T C A T C T T T A A A G T G A A C A A C T T T A coh1_a12.seq				
9786 T A G A T T T T A A T A C T T C T G T T G A C C A T C A T C T T T A A A G T G A A C A A C T T T A a909_a12.seq				
A G G T T T C A T C T G A A G C T T C T A A T G G C T T A T C A T A G T T G A C C T C T A C T T T A Majority				
9860 9870 9880 9890 9900				
9839 A G G T T T C A T C T G A A G C T T C T A A T G G C T T A T C A T A G T T G A C C T C T A C T T T coh1_a12.seq				
9836 A G G T T T C A T C T G A A G C T T C T A A T G G C T T A T C A T A G T T G A C C T C T A C T T T a909_a12.seq				
T A C T G G G G C T T G G G G T T C T G C T T C T T A C C A T T T G A C T C A A T A G T A A T G T Majority				
9910 9920 9930 9940 9950				
9889 T A C T G G G G C A G C T T G G G G T T C T G C T T C T T A C C A T T T G A C T C A A T A G T A A T G T coh1_a12.seq				
9886 T A C T G G G G C T T G G G G T T C T G C T T C T T A C C A T T T G A C T C A A T A G T A A T G T a909_a12.seq				
C A T A G A G T T T G A A G T T T T G A T T T C A C T A T C T G T T T A G C A A C T T C T G T C Majority				
9960 9970 9980 9990 10000				
9939 C A T A G A G T T T G A A G T T T T G A T T T C A C T A T C T G T T T A G C A A C T C T G T C coh1_a12.seq				
9936 C A T A G A G T T T G A A G T T T T G A T T T C A C T A T C T G T T T A G C A A C T T C T G T C a909_a12.seq				
A A T G C T T T T T C T T A T A G T C T T T A A A A G T A G C T G A A T T G T C T T T T A A T T C Majority				
10010 10020 10030 10040 10050				
9989 A A T G C T T T T T C T T A T A G T C T T T A A A A G T A G C T G A A T T G T C T T T T A A T T C coh1_a12.seq				
9986 A A T G C T T T T T C T T A T A G T C T T T A A A A G T A G C T G A A T T G T C T T T T A A T T C a909_a12.seq				
C G T C A C C T T T A A T C A G C A T T T T A G G A A T C T T A G C T T C T T T G G T C A A A G Majority				
10060 10070 10080 10090 10100				
10039 C G T C A C C T T T A A T C A G C A T T T T A G G A A T C T T A G C T T C T T T G G T C A A A G coh1_a12.seq				
10036 C G T C A C C T T T A A T C A G C A T T T T A G G A A T C T T A G C T T C T T T G G T C A A A G a909_a12.seq				
T C A C T G T T A C A G T A T A G T C T G C A C C T C T A A A C A T C A A T G G T T C T T C A C G G Majority				
10110 10120 10130 10140 10150				
10089 T C A C T G T T A C A G T A T A G T C T G C A C C T C T A A A C A T C A A T G G T T C T T C A C G G coh1_a12.seq				
10086 T C A C T G T T A C A G T A T A G T C T G C A C C T C T A A A C A T C A A T G G T T C T T C A C G G a909_a12.seq				
T A A G C A G C T T C C T C A G A A G A T G A T G T T C T G T A C A C T A G A A G C A G G A G T Majority				
10160 10170 10180 10190 10200				
10139 T A A G C A G C T T C C T C A G A A G A T G A T G T T C T G T A C A C T A G A A G C A G G A G T coh1_a12.seq				
10136 T A A G C A G C T T C C T C A G A A G A T G A T G T T C T G T A C A C T A G A A G C A G G A G T a909_a12.seq				
C T G T G G C T T G C T C T G C T C A A C A C T T G A T T G A G A A C T A G A T G T G A T G A A G Majority				
10210 10220 10230 10240 10250				
10189 C T G T G G C T T G C T C T G C T C A A C A C T T G A T T G A G A A C T A G A T G T G A T G A A G coh1_a12.seq				
10186 C T G T G G C T T G C T C T G C T C A A C A C T T G A T T G A G A A C T A G A T G T G A T G A A G a909_a12.seq				
T T A C C T G G G C T A G A A T T T T A T T T C T A A A G T A A T C C C C A C A T C A T C T G T C Majority				
10260 10270 10280 10290 10300				
10239 T T A C C T G G G C T A G A A T T T T T A T T T C T A A A G T A A T C C C C A C A T C A T C T G T C coh1_a12.seq				
10236 T T A C C T G G G C T A G A A T T T T T A T T T C T A A A G T A A T C C C C A C A T C A T C T G T C a909_a12.seq				
T T A G T T T C T T C A A C T G T T A T T G C T G G T A G A A T T A A A A A T A A G T C G T T A A Majority				
10310 10320 10330 10340 10350				
10289 T T A G T T T C T T C A A C T G T T A T T G C T G G G T A G A A T T A A A A A T A A G T C G T T A A coh1_a12.seq				
10286 T T A G T T T C T T C A A C T G T T A T T G C T G G G T A G A A T T A A A A A T A A G T C G T T A A a909_a12.seq				
A A A A G T T G T T A G G A T C A T C A A T G A C C A C A T G A T A A T T T C C A C T C T T T A G Majority				
10360 10370 10380 10390 10400				
10339 A A A A G T T G T T A G G A T C A T C A A T G A C C A C A T G A T A A T T T C C A C T C T T T A G coh1_a12.seq				
10336 A A A A G T T G T T A G G A T C A T C A A T G A C C A C A T G A T A A T T T C C A C T C T T T A G a909_a12.seq				

FIGURE 21O

GGTGTTCATCTTATGATTGATTCGATTATAAAAAGTTGACACACTCTTT Majority				
10410	10420	10430	10440	10450
10389 GCTGTTTTCTTTAATGATTGATTGACACTCTT coh1_a12.seq				
10386 GGTGTTTTCTTTAATGATTGATTGACACTCTT a909_a12.seq				
AGCATTTCGCATCCCTCCCTAACCTTAATTGATACATACATAATCTTACCTA Majority				
10460	10470	10480	10490	10500
10439 AGCATTTCGCATCCCTCCCTAACCTTAATTGATACATACATAATCTTACCTA coh1_a12.seq				
10436 AGCATTTCGCATCCCTCCCTAACCTTAATTGATACATACATAATCTTACCTA a909_a12.seq				
GAGGCCATATTCTGAAAGAAATTTACCTACAAATTGTTCTCTGAAACAA Majority				
10510	10520	10530	10540	10550
10489 GAGGCCATATTCTGAAAGAAATTTACCTACAAATTGTTCTCTGAAACAA coh1_a12.seq				
10485 GAGGCCATATTCTGAAAGAAATTTACCTACAAATTGTTCTCTGAAACAA a909_a12.seq				
TCTCCTACAGAAGTATTCTGAGAATCAATTGAAAGTTTCGGTTGTCCTC Majority				
10560	10570	10580	10590	10600
10539 TCTCCTACAGAAGTATTCTGAGAATCAATTGAAAGTTTCGGTTGTCCTC coh1_a12.seq				
10535 TCTCCTACAGAAGTATTCTGAGAATCAATTGAAAGTTTCGGTTGTCCTC a909_a12.seq				
TAATACAAAAATTCTTATCAGGTACTTGATATGGGTATTTATATTAC Majority				
10610	10620	10630	10640	10650
10589 TAATACAAAAATTCTTATCAGGTACTTGATATGGGTATTTATATTAC coh1_a12.seq				
10585 TAATACAAAAATTCTTATCAGGTACTTGATATGGGTATTTATATTAC a909_a12.seq				
TATTACCGAGTGCTTTATGAATAACATATGGTTCTTCACTTATGTTGA Majority				
10660	10670	10680	10690	10700
10639 TATTACCGAGTGCTTTATGAATAACATATGGTTCTTCACTTATGTTGA coh1_a12.seq				
10635 TATTACCGAGTGCTTTATGAATAACATATGGTTCTTCACTTATGTTGA a909_a12.seq				
TTCACTGAAACATCCCCCTTGAGAATCAATATTAACCCAGTCCTGACTC Majority				
10710	10720	10730	10740	10750
10689 TTCACTGAAACATCCCCCTTGAGAATCAATATTAACCCAGTCCTGACTC coh1_a12.seq				
10685 TTCACTGAAACATCCCCCTTGAGAATCAATATTAACCCAGTCCTGACTC a909_a12.seq				
TGCAATAACCCGGCTTGACTAGGACCTTATTATGTAGTAAACGGGACAA Majority				
10760	10770	10780	10790	10800
10739 TGCAATAACCCGGCTTGACTAGGACCTTATTATGTAGTAAACGGGACAA coh1_a12.seq				
10735 TGCAATAACCCGGCTTGACTAGGACCTTATTATGTAGTAAACGGGACAA a909_a12.seq				
CGTCCTCCAGTTAAATTGAAACCTTTACTGTAAAGACTACATCACCT Majority				
10810	10820	10830	10840	10850
10789 CGTCCTCCAGTTAAATTGAAACCTTTACTGTAAAGACTACATCACCT coh1_a12.seq				
10785 CGTCCTCCAGTTAAATTGAAACCTTTACTGTAAAGACTACATCACCT a909_a12.seq				
CCACCTTAAAGTCTTATTGAAATGTCCTGAGATTCTTAATACAGGCAA Majority				
10860	10870	10880	10890	10900
10839 CCACCTTAAAGTCTTATTGAAATGTCCTGAGATTCTTAATACAGGCAA coh1_a12.seq				
10835 CCACCTTAAAGTCTTATTGAAATGTCCTGAGATTCTTAATACAGGCAA a909_a12.seq				
CCATAAAAACCGCAATTAAATGGCTGTTGAGGCAACCCCATCAAGATGT Majority				
10910	10920	10930	10940	10950
10889 CCATAAAAACCGCAATTAAATGGCTATTGAGGCAACCCCATCAAGATGT coh1_a12.seq				
10885 CCATAAAAACCGCAATTAAATGGCTATTGAGGCAACCCCATCAAGATGT a909_a12.seq				
ATATGGTATTTTAATGACACTCCAAAAGCGTTCTGATAGTTACGGGA Majority				
10960	10970	10980	10990	11000
10939 ATATGGTATTTTAATGACACTCCAAAAGCGTTCTGATAGTTACGGGA coh1_a12.seq				
10935 ATATGGTATTTTAATGACACTCCAAAAGCGTTCTGATAGTTACGGGA a909_a12.seq				
TCCAGTTCTTGAGATAATTATCTGAACTAATCTGTCTTTCATTTGCTA Majority				
11010	11020	11030	11040	11050
10989 TCCAGTTCTTGAGATAATTATCTGAACTAATCTGTCTTTCATTTGCTA coh1_a12.seq				
10985 TCCAGTTCTTGAGATAATTATCTGAACTAATCTGTCTTTCATTTGCTA a909_a12.seq				

FIGURE 21P

127/487

Alignment Report WO 2006/078318 in method with Weighted residue weight table.

Thursday, July 29, 2004 6:49 PM

PCT/US2005/027239 Page 18

<u>C G C C T E F C T A T C H A C T A A A I F C T T A C G T T T C T A A A T A G G T A T C T G C A</u>	Majority
11060 11070 11080 11090 11100	
11039 <u>C G C C T C T C T A T C T A C T A A A T T C T T A C G T T T C T A A A T A G G T A T C T G C A</u>	coh1_a12.seq
11035 <u>C G C C T C T C T A T C T A C T A A A T T C T T A C G T T T C T A A A T A G G T A T C T G C A</u>	a909_a12.seq
<u>G C T A A C T G A G C A G T C T C A A A A A T A C C A C T A A G C A T T A A G G A A G C T T C G G C</u>	Majority
11110 11120 11130 11140 11150	
11089 <u>G C T A A C T G A G C A G T C T C A A A A A T A C C A C T A A G C A T T A A G G A A G C T T C G G C</u>	coh1_a12.seq
11085 <u>G C T A A C T G A G C A G T C T C A A A A A T A C C A C T A A G C A T T A A G G A A G C T T C G G C</u>	a909_a12.seq
<u>A A T A G A A C C A G C T T G C T A A T T T G A T T T T T A T C A T C T A G C G C T T C T T</u>	Majority
11160 11170 11180 11190 11200	
11139 <u>A A T A G A A C C A G C T T G C T A A T T T G A T T T T T A T C A T C T A G C G C T T C T T</u>	coh1_a12.seq
11135 <u>A A T A G A A C C A G C T T G C T A A T T T G A T T T T T A T C A T C T A G C G C T T C T T</u>	a909_a12.seq
<u>T A A G T T G C T G A A T C T C T T C T G T T T T C A A T A A G A A G T T G C T G T T C T</u>	Majority
111210 111220 111230 111240 111250	
11189 <u>T A A G T T G C T G A A T C T C T T C T G T T T T C A A T A A G A A G T T G C T G T T C T</u>	coh1_a12.seq
11185 <u>T A A G T T G C T G A A T C T C T T C T G T T T T C A A T A A G A A G T T G C T G T T C T</u>	a909_a12.seq
<u>A A C A T A A T T C T A G C A A G T C T T T C T T T T A A T T T T T A A T C T T C C A T</u>	Majority
111260 111270 111280 111290 111300	
111239 <u>A A C A T A A T T C T A G C A A G T C T T T C T T T T A A T T T T T A A T C T T C C A T</u>	coh1_a12.seq
111235 <u>A A C A T A A T T C T A G C A A G T C T T T C T T T T A A T T T T T A A T C T T C C A T</u>	a909_a12.seq
<u>C G C G A T T A C T T C C T T A A C T G A A C C T T A A A T T A T C G T T T A G A T A T T A T A C</u>	Majority
111310 111320 111330 111340 111350	
111289 <u>C G C G A T T A C T T C C T T A A C T G A A C C T T A A A T T A T C G T T T A G A T A T T A T A C</u>	coh1_a12.seq
111285 <u>C G C G A T T A C T T C C T T A A C T G A A C C T T A A A T T A T C G T T T A G A T A T T A T A C</u>	a909_a12.seq
<u>A A A G T T C T A A C C T T T A A A C T C A T T T T T G T C C T G T T T T T C T C A A A A A</u>	Majority
111360 111370 111380 111390 111400	
111339 <u>A A A G T T C T A A C C T T T A A A C T C A T T T T T G T C C T G T T T T T C T C A A A A A</u>	coh1_a12.seq
111335 <u>A A A G T T C T A A C C T T T A A A C T C A T T T T T G T C C T G T T T T T C T C A A A A A</u>	a909_a12.seq
<u>A G T C T A T G C T A A A T T A A C A T T T T G A T A A T T T T T G A A A A A T C T C A T C G A</u>	Majority
111410 111420 111430 111440 111450	
111389 <u>A G T C T A T G C T A A A T T A A C A T T T T G A T A A T T T T T G A A A A A T C T C A T C G A</u>	coh1_a12.seq
111385 <u>A G T C T A T G C T A A A T T A A C A T T T T G A T A A T T T T T G A A A A A T C T C A T C G A</u>	a909_a12.seq
<u>A G T C A T T T C T T T G A A A G C T C G A A T T C T A G G C A T T A A A A G C C A T A T A</u>	Majority
111460 111470 111480 111490 111500	
111439 <u>A G T C A T T T C T T T G A A A G C T C G A A T T C T A G G C A T T A A A A G C C A T A T A</u>	coh1_a12.seq
111435 <u>A G T C A T T T C T T T G A A A G C T C G A A T T C T A G G C A T T A A A A G C C A T A T A</u>	a909_a12.seq
<u>T C A A A T T G A T A T G G C T T T T T A T T A A A A C A A A A A G A A T C A A T A G</u>	Majority
111510 111520 111530 111540 111550	
111489 <u>T C A A A T T G A T A T G G C T T T T T A T T A A A A C A A A A A G A A T C A A T A G</u>	coh1_a12.seq
111476 <u>T C A A A T T G A T A T G G C T T T T T A T T A A A A C A A A A A G A A T C A A T A G</u>	a909_a12.seq
<u>G A C A A T A G C C G T C A A T T A G T G A C A T A A T C T A T T A C A G A T T A A G T T C T T A T</u>	Majority
111560 111570 111580 111590 111600	
111539 <u>G A C A A T A G C C G T C A A T T A G T G A C A T A A T C T A T T A C A G A T T A A G T T C T T T</u>	coh1_a12.seq
111492 <u>G A C A A T A G C C G T C A A T T A G T G A C A T A A T C T A T T A C A G A T T A A G T T C T T T</u>	a909_a12.seq
<u>T G A A T A A T A A T C C A A C T T T C A A C T G T T T T T C C A T G T G A A A T G T T C</u>	Majority
111610 111620 111630 111640 111650	
111589 <u>T G A A T A A T A A T C C A A C T T T C A A C T G T T T T T C C A T G T G A A A T G T T C</u>	coh1_a12.seq
111542 <u>T G A A T A A T A A T C C A A C T T T C A A C T G T T T T T C C A T G T G A A A T G T T C</u>	a909_a12.seq
<u>T T T A A T T C T T T T A G C A A T A T T C T G T T G T A G T T T C T C T T A A T G C C T T A T</u>	Majority
111660 111670 111680 111690 111700	
111639 <u>T T T A A T T C T T T T A G C A A T A T T C T G T T G T A G T T T C T C T T A A T G C C T T A T</u>	coh1_a12.seq
111592 <u>T T T A A T T C T T T T A G C A A T A T T C T G T T G T A T T T C T C T T A A T G C C T T A T</u>	a909_a12.seq

FIGURE 21Q

C T T T A G T A T A A A T C A A G A G A T T C A T G G A G T G A C T G A G T A T T T C T T C C Majority

11710 11720 11730 11740 11750

11689 C T T T A C T A A T A A A T C A A G A G A T T C A T G G A G T G A C T G A G T A T T T C T T C C cohi_a12.seq

11642 C T T T A C T A A T A A A T C A A G A G A T T C A T G G A G T G A C T G A G T A T T T C T T C C a909_a12.seq

A T G A T G A T T C C T A A C T C A G G G C T A T C A A T A A C T C A A C T G T T C C A C C C C G Majority

11760 11770 11780 11790 11800

11739 A T G A T G A T T C C T A A C T C A G G G C T A T C A A T A A C T C A A C T G T T C C A C C C C G cohi_a12.seq

11692 A T G A T G A T T C C T A A C T C A G G G C T A T C A A T A A C T C A A C T G T T C C A C C C C G a909_a12.seq

A T C T G T T G C A A T A A T A G C A C T C G A A A G T A G A C C A G C T T C T A A A A T A G A G G Majority

11810 11820 11830 11840 11850

11789 A T C T G T T G C A A T A A T A G C A C T C G A A A G T A G A C C A G C T T C T A A A A T A G A G G cohi_a12.seq

11742 A T C T G T T G C A A T A A T A G C A C T C G A A A G T A G A C C A G C T T C T A A A A T A G A G G a909_a12.seq

T T G G T A A T C C C T C T G G A T A C A T T G A T G G G T A A A C A A A G A T A T C T G T C T G T Majority

11860 11870 11880 11890 11900

11839 T T G G T A A T C C C T C T G G A T A C A T T G A T G G G T A A A C A A A G A T A T C T G T C T G T cohi_a12.seq

11792 T T G G T A A T C C C T C T G G A T A C A T T G A T G G G T A A A C A A A G A T A T C T G T C T G T a909_a12.seq

G C C A T T A A A G A C A T A G T C T G T C A A A G T T T A A T T C C C C A A A A G T T A A T Majority

11910 11920 11930 11940 11950

11889 G C C A T T A A A G A C A T A G T C T G T C A A A G T T T A A T T C C C C A A A A G T T A A T cohi_a12.seq

11842 G C C A T T A A A G A C A T A G T C T G T C A A A G T T T A A T T C C C C A A A A G T T A A T a909_a12.seq

C T G T T T G G A C T G A T A T T T C T C T T C A A A T G T G C T A A T T C A G G T C C G T C T C Majority

11960 11970 11980 11990 12000

11839 C T G T T T G G A C T G A T A T T T C T C T T C A A A T G T G C T A A T T C A G G T C C G T C T C cohi_a12.seq

11892 C T G T T T G G A C T G A T A T T T C T C T T C A A A T G T G C T A A T T C A G G T C C G T C T C a909_a12.seq

C T G C A A T C T G T A A A A T A A C A T T T C A G A G T A C T G T G A C A T C G A A A A T G C T Majority

12010 12020 12030 12040 12050

11989 C T G C A A T C T G T A A A A T A A C A T T T C A G A G T A C T G T G A C A T C G A A A A T G C T cohi_a12.seq

11942 C T G C A A T C T G T A A A A T A A C A T T T C A G A G T A C T G T G A C A T C G A A A A T G C T a909_a12.seq

T C T A A G A G C A A T T C A A T G C C T T T T C T T A A T A A T T C T A C C A G G C A T A A G T Majority

12060 12070 12080 12090 12100

12039 T C T A A G A G C A A T T C A A T G C C T T T T C T T A A T A A T T C T A C C A G G C A T A A G T cohi_a12.seq

11992 T C T A A G A G C A A T T C A A T G C C T T T T C T T A A T A A T T C T A C C A G G C A T A A G T a909_a12.seq

G A T G A A A A T A T C A T C A G C A G A T T T T C A A G G T A A G G C G T G T C A G C A A A A T Majority

12110 12120 12130 12140 12150

12089 G A T G A A A A T A T C A T C A G C A G A T T T T C A A G G T A A G G C G T G T C A G C A A A A T cohi_a12.seq

12042 G A T G A A A A T A T C A T C A G C A G A T T T T C A A G G T A A G G C G T G T C A G C A A A A T a909_a12.seq

C A G A G G C C T A G A C T T C A G A T A C C G A A T T T A A A T A A C T C C T T A G C T T G C T Majority

12160 12170 12180 12190 12200

12139 C A G A G C C T A G A C T T C A G A T A C C G A A T T T A A A T A A C T C C T T T A G C T T G C T cohi_a12.seq

12092 C A G A G G C C T A G A C T T C A G A T A C C G A A T T T A A A T A A C T C C T T T A G C T T G C T a909_a12.seq

A T A T T A A A A T G T T T A A C C A T C A A C G C T T C T C T T G G A T A C C G C A T A A A A Majority

12210 12220 12230 12240 12250

12189 A T A T T A A A A T G T T T A A C C A T C A A C G C T T C T C T T G G A T A C C G C A T A A A A cohi_a12.seq

12142 A T A T T A A A A T G T T T A A C C A T C A A C G C T T C T C T T G G A T A C C G C A T A A A A a909_a12.seq

A T C T G G A C C G A T A G T G C T T A A C A C G G C G T G T G A G A A G A T G T T C A T A G A T A G Majority

12260 12270 12280 12290 12300

12239 A T C T G G A C C G A T A G T G C T T A A C A C G G C G T G T G A G A A G A T G T T C A T A G A T A G cohi_a12.seq

12192 A T C T G G A C C G A T A G T G C T T A A C A C G G C G T G T G A G A A G A T G T T C A T A G A T A G a909_a12.seq

C T C C A A A G A A A A T C T A A A A A C G A T T T G A C A G A A A A A T G A C T T G A C C C A Majority

12310 12320 12330 12340 12350

12289 C T C C A A A G A A A A T C T A A A A A C G A T T T G A C A G A A A A A T G A C T T G A C C C A cohi_a12.seq

12242 C T C C A A A G A A A A T C T A A A A A C G A T T T G A C A G A A A A A T G A C T T G A C C C A a909_a12.seq

FIGURE 21R

129/487

Alignment Report of WO 2006/078318 in motion with Weighted residue weight table.
Thursday, July 29, 2009

PCT/US2005/027239 20

T G G T C T A A A A C A A T A C T A G G T A A A C G G C T G C T T C T T G C A A A A G A T A G C C C Majority				
12360 12370 12380 12390 12400				
12339 T G G T C T A A A A C A A T A C T A G G T A A A C G G G T G C T T C T T G C A A A A G A T A G C C C coh1_a12.seq				
12292 T G G T C T A A A A C A A T A C T A G G T A A A C G G G T G C T T C T T G C A A A A G A T A G C C C a909_a12.seq				
<u>T T C T A G C G T T G T T A A C T G A A A A C G T G T A T T A C A A A T C A C A A A A T C A A T A T</u> Majority				
12410 12420 12430 12440 12450				
12389 T T C T A G C G T T G T T A A C T G A A A A C G T G T A T T A C A A A T C A C A A A A T C A A T A coh1_a12.seq				
12342 T T C T A G C G T T G T T A A C T G A A A A C G T G T A T T A C A A A T C A C A A A A T C A A T A a909_a12.seq				
<u>T T T C A T C T G A A A C A T G T T C A T C A G G C T G T T G T A T T C T C G A T T T T G T T A</u> Majority				
12460 12470 12480 12490 12500				
12439 T T T C A T C T G A A A C A T G T T C A T C A G G C T G T T G T A T T C T C G A T T T T G T T A coh1_a12.seq				
12392 T T T C A T C T G A A A C A T G T T C A T C A G G C T G T T G T A T T C T C G A T T T T G T T A a909_a12.seq				
<u>A T A A T A G G A T A G C C C T G C T T G A C A A T G T T T T G G T C G G T A A A C G G T A A A A T</u> Majority				
12510 12520 12530 12540 12550				
12489 A T A A T A G G A T A G C C C T G C T T G A C A A T T T T T G G T C G G T A A A C G G T A A A A T coh1_a12.seq				
12442 A T A A T A G G A T A G C C C T G C T T G A C A A T G T T T T G G T C G G T A A A C G G T A A A A T a909_a12.seq				
<u>T T T T C T A C C C C T T G T C T T C A T C T A T A A T C G G T A A A T C A T C A T G A T T A G T T G</u> Majority				
12560 12570 12580 12590 12600				
12539 T T T T C T A C C C C T T G T C T T C A T C T A T A A T C G G T A A A T C A T C A T G A T T A G T T G coh1_a12.seq				
12492 T T T T C T A C C C C T T G T C T T C A T C T A T A A T C G G T A A A T C A T C A T G A T T A G T T G a909_a12.seq				
<u>T T A C A A T A A C A A C A C G G T A G C C A C C G C T T A A C C A A A T C T G C T G T C A T T T A</u> Majority				
12610 12620 12630 12640 12650				
12589 T T A C A A T A A C A A C A C G G T A G C C A C C G C T T A A C C A A A T C T G C T G T C A T T T A coh1_a12.seq				
12542 T T A C A A T A A C A A C A C G G T A G C C A C C G C T T A A C C A A A T C T G C T G T C A T T T A a909_a12.seq				
<u>T C T G T A T A A C G T T C A A T A C C T C C G A G G A A G G G T A G A T A A T A T C T G A G A A A</u> Majority				
12660 12670 12680 12690 12700				
12639 T C T G T A T A A C G T T C A A T A C C T C C G A G G A A G G G T A G A T A A T A T C T G A G A A coh1_a12.seq				
12592 T C T G T A T A A C G T T C A A T A C C T C C G A G G A A G G G T A G A T A A T A T C T G A G A A a909_a12.seq				
<u>A A C A G G C A A C T G T T T T A C C T T A T T T C A T A T T A T C C A C T T I C A T C A A T</u> Majority				
12710 12720 12730 12740 12750				
12689 A A C A G G C A A C T G T T T T A C C T T A T T T C A T A T T A T C C A C T T I C A T C A A T coh1_a12.seq				
12642 A A C A G G C A A C T G T T T T A C C T T A T T T C A T A T T A T C C A C T T I C A T C A A T a909_a12.seq				
<u>A A G C C A T C T T T A A G C C T T A A T C A T A G C A A C T A A T T T T G C T C T T T G</u> Majority				
12760 12770 12780 12790 12800				
12739 A A G C C A T C T T T A A G C C T T A A T C A T A G C A A C T A A T T T T G C T C T T T G coh1_a12.seq				
12692 A A G C C A T C T T T A A G C C T T A A T C A T A G C A A C T A A T T T T G C T C T T T G a909_a12.seq				
<u>C T C T T C T G C T A C C A A C A C T C G A A C A A A T T C A T T I C G G C A T A A A T A C T A A A T</u> Majority				
12810 12820 12830 12840 12850				
12789 C T C T T C T G C T A C C A A C A C T C G A A C A A A T T C A T T I C G G C A T A A A T A C T A A A T coh1_a12.seq				
12742 C T C T T C T G C T A C C A A C A C T C G A A C A A A T T C A T T I C G G C A T A A A T A C T A A A T a909_a12.seq				
<u>A T T T G T G C C C G T T C T T C T A C C A T A T T T T T A T A A T A A A T C G C A T T G</u> Majority				
12860 12870 12880 12890 12900				
12839 A T T T G T G C C C G T T C T T C T A C C A T A T T T T T A T A A T A A A T C G C A T T G coh1_a12.seq				
12792 A T T T G T G C C C G T T C T T C T A C C A T A T T T T T A T A A T A A A T C G C A T T G a909_a12.seq				
<u>C G T A T C A T G T A A T A T T T C G A A A T G G T G A A T G A T T C A A T A C A T G A A A A A C</u> Majority				
12910 12920 12930 12940 12950				
12889 C G T A T C A T G T A A T A T T T C G A A A T G G T G A A T G A T T C A A T A C A T G A A A A A C coh1_a12.seq				
12842 C G T A T C A T G T A A T A T T T C G A A A T G G T G A A T G A T T C A A T A C A T G A A A A A C a909_a12.seq				
<u>A T G G C C A A A T T T T A A C T C G T G A A G A G T G T C C A A T T C G T G T A A A A G A C</u> Majority				
12960 12970 12980 12990 13000				
12939 A T G G C C A A A T T T T T A A C T C G T G A A G A G T G T C C A A T T C G T G T A A A A G A C coh1_a12.seq				
12892 A T G G C C A A A T T T T T A A C T C G T G A A G A G T G T C C A A T T C G T G T A A A A G A C a909_a12.seq				

FIGURE 21S

CAATAAAATTAACCTGATAAGTCTTATATCCCACATCTCTGACAGACGATAA Majority
 13010 13020 13030 13040 13050

12989 CAATAAAATTAACCTGATAAGTCTTATATCCCACATCTCTGACAGACGATAA coh1_a12.seq
 12942 CAATAAAATTAACCTGATAAGTCTTATATCCCACATCTCTGACAGACGATAA a909_a12.seq

TTCATTTCAGAGTCAACAAAAATCAATAAACATCTCTCTGCAAAGCCAGA Majority
 13060 13070 13080 13090 13100

13039 TTCATTTCAGAGTCAACAAAAATCAATAAACATCTCTCTGCAAAGCCAGA coh1_a12.seq
 12992 TTCATTTCAGAGTCAACAAAAATCAATAAACATCTCTCTGCAAAGCCAGA a909_a12.seq

TGTTCCTCAAAAAACGCTCGTTTCATTAAGCAGCCGAAGTAATACACT Majority
 13110 13120 13130 13140 13150

13089 TGTTCCTCAAAAAACGCTCGTTTCATTAAGCAGCCGAAGTAATACACT coh1_a12.seq
 13042 TGTTCCTCAAAAAACGCTCGTTTCATTAAGCAGCCGAAGTAATACACT a909_a12.seq

CTTCAAATTCTTTATAGTCAAATTCTGCATCACTAAATCTTCACGGTT Majority
 13160 13170 13180 13190 13200

13139 CTTCAAATTCTTTATAGTCAAATTCTGCATCACTAAATCTTCACGGTT coh1_a12.seq
 13092 CTTCAAATTCTTTATAGTCAAATTCTGCATCACTAAATCTTCACGGTT a909_a12.seq

ATATCTTGATACAAACAAAGATAACATACCGACCTTAGGTAATGAAGGT Majority
 13210 13220 13230 13240 13250

13189 ATATCTTGATACAAACAAAGATAACATACCGACCTTAGGTAATGAAGGT coh1_a12.seq
 13142 ATATCTTGATACAAACAAAGATAACATACCGACCTTAGGTAATGAAGGT a909_a12.seq

ATTTTCATAATTATCTATCAAAATCACCTAGGACAACCGAACATCTGATCTA Majority
 13260 13270 13280 13290 13300

13239 ATTTTCATAATTATCTATCAAAATCACCTAGGACAACCGAACATCTGATCTA coh1_a12.seq
 13192 ATTTTCATAATTATCTATCAAAATCACCTAGGACAACCGAACATCTGATCTA a909_a12.seq

AAGTCAAGAACCAATCAAATTCTTGCTACTGCAAATTGACCGATACAG Majority
 13310 13320 13330 13340 13350

13289 AAGTCAAGAACCAATCAAATTCTTGCTACTGCAAATTGACCGATACAG coh1_a12.seq
 13242 AAGTCAAGAACCAATCAAATTCTTGCTACTGCAAATTGACCGATACAG a909_a12.seq

TTCAAAAGCATATGCCAAATTCTTTATTTCTGTAAATAATCAACAGTTAG Majority
 13360 13370 13380 13390 13400

13339 TTCAAAAGCATATGCCAAATTCTTTATTTCTGTAAATAATCAACAGTTAG coh1_a12.seq
 13292 TTCAAAAGCATATGCCAAATTCTTTATTTCTGTAAATAATCAACAGTTAG a909_a12.seq

GTGCCCTCTCATTATAATCGGCCACTAATTGAGAAATTCTTCCTTAT Majority
 13410 13420 13430 13440 13450

13389 GTGCCCTCTCATTATAATCGGCCACTAATTGAGAAATTCTTCCTTAT coh1_a12.seq
 13342 GTGCCCTCTCATTATAATCGGCCACTAATTGAGAAATTCTTCCTTAT a909_a12.seq

TTTCGAGCCATTATCTACGATGTAGATATGGCTTACTTGAGGATAAATT Majority
 13460 13470 13480 13490 13500

13439 TTTCGAGCCATTATCTACGATGTAGATATGGCTTACTTGAGGATAAATT coh1_a12.seq
 13392 TTTCGAGCCATTATCTACGATGTAGATATGGCTTACTTGAGGATAAATT a909_a12.seq

GCTCGAATGTCGTGATGTAAGCGTTCAATATTGGGGTTAAAGGTGACAAAT Majority
 13510 13520 13530 13540 13550

13489 GCTCGAATGTCGTGATGTAAGCGTTCAATATTGGGGTTAAAGGTGACAAAT coh1_a12.seq
 13442 GCTCGAATGTCGTGATGTAAGCGTTCAATATTGGGGTTAAAGGTGACAAAT a909_a12.seq

ACCCGCTAAATATTTCATGTTCTATGCTCTTTCTAAATCTCTAAATAA Majority
 13560 13570 13580 13590 13600

13539 ACCCGCTAAATATTTCATGTTCTATGCTCTTTCTAAATCTCTAAATAA coh1_a12.seq
 13492 ACCCGCTAAATATTTCATGTTCTATGCTCTTTCTAAATCTCTAAATAA a909_a12.seq

CTGAATGACTGGTGCTTGGTTATAAAACGATACCGACATAGATAGTTA Majority
 13610 13620 13630 13640 13650

13589 CTGAATGACTGGTGCTTGGTTATAAAACGATACCGACATAGATAGTTA coh1_a12.seq
 13542 CTGAATGACTGGTGCTTGGTTATAAAACGATACCGACATAGATAGTTA a909_a12.seq

FIGURE 21T

131/487

Alignment Report WO 2006/078318 in method with Weighted residue weight table.
Thursday, July 29, 2004 6:49 PMPCT/US2005/027239
Page 22

<u>C T G C T A G C T A A A G C T C G A A T T G A C A T T A A T T A C C A A T G A T A C T G A C A T T G A</u> Majority				
13660	13670	13680	13690	13700
13639 <u>C T G C T A C T A A A C T T T G A A T G A C A T A A T T A C C A A T G A T A C T G A C A T T G A</u> coh1_a12.seq				
13592 <u>C T G C T A C T A A A C T T T G A A T G A C A T A A T T A C C A A T G A T A C T G A C A T T G A</u> a909_a12.seq				
<u>G T A T T G A T A T A A T A G A G T A C A G C T C C A C T A A G A G T A G C A G G C A A T T A A A T A</u> Majority				
13710	13720	13730	13740	13750
13689 <u>G T A T T G A T A T A A T A G A G T A C A G C T C C A C T A A G A G T A G C A G G C A A T T A A A T A</u> coh1_a12.seq				
13642 <u>G T A T T G A T A T A A T A G A G T A C A G C T C C A C T A A G A G T A G C A G G C A A T T A A A T A</u> a909_a12.seq				
<u>G C C G A G C A T T C C T C T G T T A A T T C T T A A A A G T A A A T A C A T C T C T T A A A G</u> Majority				
13760	13770	13780	13790	13800
13739 <u>G C C G A G C A T T C C T C T G T T A A T T C T T A A A A G T A A A T A C A T C T C T T A A A G</u> coh1_a12.seq				
13692 <u>G C C G A G C A T T C C T C T G T T A A T T C T T A A A A G T A A A T A C A T C T C T T A A A G</u> a909_a12.seq				
<u>A G A T A G C T T G A T A T A G G G A G A C A A T A A A T T C A G T A A T A A C T G T A G A G A T A</u> Majority				
13810	13820	13830	13840	13850
13789 <u>A G A T A G C T T G A T A T A G G G A G A C A A T A A A T T C A G T A A T A A C T G T A G A G A T A</u> coh1_a12.seq				
13742 <u>A G A T A G C T T G A T A T A G G G A G A C A A T A A A T T C A G T A A T A A C T G T A G A G A T A</u> a909_a12.seq				
<u>A T A G C T C C C A T A G C A C C T A A A A T T G G T A T T A A A A G T A T A T A A G C A C A A C</u> Majority				
13860	13870	13880	13890	13900
13839 <u>A T A G C T C C C A T A G C A C C T A A A A T T G G T A T T A A A A G T A T A T A A G C A C A A C</u> coh1_a12.seq				
13792 <u>A T A G C T C C C A T A G C A C C T A A A A T T G G T A T T A A A A G T A T A T A A G C A C A A C</u> a909_a12.seq				
<u>A T T T G C C A C A A G T C C A A T A A C T G C A G A C A T T G T G T A A G C T T T T G T A C G T C</u> Majority				
13910	13920	13930	13940	13950
13889 <u>A T T T G C C A C A A G T C C A A T A A C T G C A G A C A T T G T G T A A G C T T T T G T A C G T C</u> coh1_a12.seq				
13842 <u>A T T T G C C A C A A G T C C A A T A A C T G C A G A C A T T G T G T A A G C T T T T G T A C G T C</u> a909_a12.seq				
<u>T T G A A G C C A G T A G A T A C T G T G T C C C T A A A G C G T T A C C A T A A G A A A T G C A A</u> Majority				
13960	13970	13980	13990	14000
13939 <u>T T G A A G C C A G T A G A T A C T G T G T C C C T A A A G C G T T A C C A T A A G A A A T G C A A</u> coh1_a12.seq				
13892 <u>T T G A A G C C A G T A G A T A C T G T G T C C C T A A A G C G T T A C C A T A A G A A A T G C A A</u> a909_a12.seq				
<u>A T G A T C A T C A A A</u> Majority				
14010				
13989 <u>A T G A T C A T C A A A</u> coh1_a12.seq				
13942 <u>A T G A T C A T C A A A</u> a909_a12.seq				

Decoration 'Decoration #1': Shade (with solid black) residues that differ from the consensus.

FIGURE 21U

Alignment Report of gbs: WO 2006/078318 method with PAM250 residue weight table.
 Thursday, July 29, 2004 6:57 PM

	410	420	430	440	450	Majority
101	F D L L A S D G T A V K W T D A L I K A N T N K N Y I A G E A V T G Q P I K L K S H T D G T F E I K sag645_2603.pep					
101	F D L L A S D G T A V K W T D A L I K A N T N K N Y I A G E A V T G Q P I K L K S H T D G T F E I K sag645_a909.pep					
101	F D L L A S D G T A V K W T D A L I K A N T N K N Y I A G E A V T G Q P I K L K S H T D G T F E I K sag645_cjb111.pep					
101	F D L L A S D G T A V K W T D A L I K A N T N K N Y I A G E A V T G Q P I K L K S H T D G T F E I K sag645_coh1.pep					
101	F D L L A S D G T A V K W T D A L I K A N T N K N Y I A G E A V T G Q P I K L K S H T D G T F E I K sag645_nem316.pep					
	460	470	480	490	500	Majority
151	G L A Y A V D A N A E G T A V T Y K L K E T K A P E G Y V I P D K E I E F T V S Q T S Y N T K P T D sag645_2603.pep					
151	G L A Y A V D A N A E G T A V T Y K L K E T K A P E G Y V I P D K E I E F T V S Q T S Y N T K P T D sag645_a909.pep					
151	G L A Y A V D A N A E G T A V T Y K L K E T K A P E G Y V I P D K E I E F T V S Q T S Y N T K P T D sag645_cjb111.pep					
51	G L A Y A V D A N A E G T A V T Y K L K E T K A P E G Y V I P D K E I E F T V S Q T S Y N T K P T D sag645_coh1.pep					
51	G L A Y A V D A N A E G T A V T Y K L K E T K A P E G Y V I P D K E I E F T V S Q T S Y N T K P T D sag645_nem316.pep					
	510	520	530	540	550	Majority
01	I T V D S A D A T P D T I K N N K R P S I P N T G G I C T A I F V A I G A A V M A F A V K G M K R R sag645_2603.pep					
01	I T V D S A D A T P D T I K N N K R P S I P N T G G I C T A I F V A I G A A V M A F A V K G M K R R sag645_a909.pep					
01	I T V D S A D A T P D T I K N N K R P S I P N T G G I C T A I F V A I G A A V M A F A V K G M K R R sag645_cjb111.pep					
01	I T V D S A D A T P D T I K N N K R P S I P N T G G I C T A I F V A I G A A V M A F A V K G M K R R sag645_coh1.pep					
01	I T V D S A D A T P D T I K N N K R P S I P N T G G I C T A I F V A I G A A V M A F A V K G M K R R sag645_nem316.pep					
T K D N						
						Majority
51	T K D N sag645_2603.pep					
51	T K D N sag645_a909.pep					
51	T K D N sag645_cjb111.pep					
51	T K D N sag645_coh1.pep					
51	T K D N sag645_nem316.pep					

ecoration 'Decoration #1': Shade (with solid black) residues that differ from the Consensus.

FIGURE 22A

134/487

Alignment Report WO 2006/078318 al method with PAM250 residue weight table.
Thursday, July 29, 2004 7:05 PMPCT/US2005/027239
Page 1

					Majority
	10	20	30	40	50
1	M K K R Q K I W R G L S V T L L E T S Q I P F G I L V Q G E T Q D T N Q A L G K V I V K K T G D N A				sag649_2603.pep
1	M K K R Q K I W R G L S V T L L I L S Q I P F G I L V Q G E T Q D T N Q A L G K V I V K K T G D N A				sag649_coh1.pep
1	M K K R Q K I W R G L S V T L L I L S Q I P F G I L V Q G E T Q D T N Q A L G K V I V K K T G D N A				sag649_cjb111.pep
1	M K K R Q K I W R G L S V T L L I L S Q I P F G I L V Q G E T Q D T N Q A L G K V I V K K T G D N A				sag649_18rs21.pep
1	M K K R Q K I W R G L S V T L L I L S Q I P F G I L V Q G E T Q D T N Q A L G K V I V K K T G D N A				sag649_nem316.pep
	T P L G K A T F V L K N D N D K S E T S H E T V E G S G E A T F E N I K P G D Y T L R E E T A P I G				Majority
	60	70	80	90	100
51	T P L G K A T F V L K N D N D K S E T S H E T V E G S G E A T F E N I K P G D Y T L R E E T A P I G				sag649_2603.pep
51	T P L G K A T F V L K N D N D K S E T S H E T V E G S G E A T F E N I K P G D Y T L R E E T A P I G				sag649_coh1.pep
51	T P L G K A T F V L K N D N D K S E T S H E T V E G S G E A T F E N I K P G D Y T L R E E T A P I G				sag649_cjb111.pep
51	T P L G K A T F V L K N D N D K S E T S H E T V E G S G E A T F E N I K P G D Y T L R E E T A P I G				sag649_18rs21.pep
51	T P L G K A T F V L K N D N D K S E T S H E T V E G S G E A T F E N I K P G D Y T L R E E T A P I G				sag649_nem316.pep
	Y K K T D K T W K V K V A D N G A T I I E G M D A D K A E K R K E V L N A Q Y P K S A I Y E D T K E				Majority
	110	120	130	140	150
101	Y K K T D K T W K V K V A D N G A T I I E G M D A D K A E K R K E V L N A Q Y P K S A I Y E D T K E				sag649_2603.pep
101	Y K K T D K T W K V K V A D N G A T I I E G M D A D K A E K R K E V L N A Q Y P K S A I Y E D T K E				sag649_coh1.pep
101	Y K K T D K T W K V K V A D N G A T I I E G M D A D K A E K R K E V L N A Q Y P K S A I Y E D T K E				sag649_cjb111.pep
101	Y K K T D K T W K V K V A D N G A T I I E G M D A D K A E K R K E V L N A Q Y P K S A I Y E D T K E				sag649_18rs21.pep
101	Y K K T D K T W K V K V A D N G A T I I E G M D A D K A E K R K E V L N A Q Y P K S A I Y E D T K E				sag649_nem316.pep
	N Y P L V N V E G S K V G E Q Y K A L N P I N G K D G R R E I A E G W L S K I T G V N D L D K N K				Majority
	160	170	180	190	200
151	N Y P L V N V E G S K V G E Q Y K A L N P I N G K D G R R E I A E G W L S K I T G V N D L D K N K				sag649_2603.pep
151	N Y P L V N V E G S K V G E Q Y K A L N P I N G K D G R R E I A E G W L S K I T G V N D L D K N K				sag649_coh1.pep
151	N Y P L V N V E G S K V G E Q Y K A L N P I N G K D G R R E I A E G W L S K I T G V N D L D K N K				sag649_cjb111.pep
151	N Y P L V N V E G S K V G E Q Y K A L N P I N G K D G R R E I A E G W L S K I T G V N D L D K N K				sag649_18rs21.pep
151	N Y P L V N V E G S K V G E Q Y K A L N P I N G K D G R R E I A E G W L S K I T G V N D L D K N K				sag649_nem316.pep
	Y K I E L T V E G K T T V E T K E L N Q P L D V V V L L D N S N S M N N E R A N N S Q R A L K A G E				Majority
	210	220	230	240	250
201	Y K I E L T V E G K T T V E T K E L N Q P L D V V V L L D N S N S M N N E R A N N S Q R A L K A G E				sag649_2603.pep
201	Y K I E L T V E G K T T V E T K E L N Q P L D V V V L L D N S N S M N N E R A N N S Q R A L K A G E				sag649_coh1.pep
201	Y K I E L T V E G K T T V E T K E L N Q P L D V V V L L D N S N S M N N E R A N N S Q R A L K A G E				sag649_cjb111.pep
201	Y K I E L T V E G K T T V E T K E L N Q P L D V V V L L D N S N S M N N E R A N N S Q R A L K A G E				sag649_18rs21.pep
201	Y K I E L T V E G K T T V E T K E L N Q P L D V V V L L D N S N S M N N E R A N N S Q R A L K A G E				sag649_nem316.pep
	A V E K L I D K I T S N K D N R V A L V T Y A S T I F D G T E A T V S K G V A D Q N G K A L N D S V				Majority
	260	270	280	290	300
251	A V E K L I D K I T S N K D N R V A L V T Y A S T I F D G T E A T V S K G V A D Q N G K A L N D S V				sag649_2603.pep
251	A V E K L I D K I T S N K D N R V A L V T Y A S T I F D G T E A T V S K G V A D Q N G K A L N D S V				sag649_coh1.pep
251	A V E K L I D K I T S N K D N R V A L V T Y A S T I F D G T E A T V S K G V A D Q N G K A L N D S V				sag649_cjb111.pep
251	A V E K L I D K I T S N K D N R V A L V T Y A S T I F D G T E A T V S K G V A D Q N G K A L N D S V				sag649_18rs21.pep
251	A V E K L I D K I T S N K D N R V A L V T Y A S T I F D G T E A T V S K G V A D Q N G K A L N D S V				sag649_nem316.pep
	S W D Y H K T T F T A T T H N Y S Y L N L T N D A N E V N I L K S R I P K E A E H I N G D R T L Y Q				Majority
	310	320	330	340	350
301	S W D Y H K T T F T A T T H N Y S Y L N L T N D A N E V N I L K S R I P K E A E H I N G D R T L Y Q				sag649_2603.pep
301	S W D Y H K T T F T A T T H N Y S Y L N L T N D A N E V N I L K S R I P K E A E H I N G D R T L Y Q				sag649_coh1.pep
301	S W D Y H K T T F T A T T H N Y S Y L N L T N D A N E V N I L K S R I P K E A E H I N G D R T L Y Q				sag649_cjb111.pep
301	S W D Y H K T T F T A T T H N Y S Y L N L T N D A N E V N I L K S R I P K E A E H I N G D R T L Y Q				sag649_18rs21.pep
301	S W D Y H K T T F T A T T H N Y S Y L N L T N D A N E V N I L K S R I P K E A E H I N G D R T L Y Q				sag649_nem316.pep
	F G A T F T Q K A L M K A N E I L E T Q S S N A R K K L I F H V T D G V P T H S Y A I N F N P Y I S				Majority
	360	370	380	390	400
351	F G A T F T Q K A L M K A N E I L E T Q S S N A R K K L I F H V T D G V P T H S Y A I N F N P Y I S				sag649_2603.pep
351	F G A T F T Q K A L M K A N E I L E T Q S S N A R K K L I F H V T D G V P T H S Y A I N F N P Y I S				sag649_coh1.pep
351	F G A T F T Q K A L M K A N E I L E T Q S S N A R K K L I F H V T D G V P T H S Y A I N F N P Y I S				sag649_cjb111.pep
351	F G A T F T Q K A L M K A N E I L E T Q S S N A R K K L I F H V T D G V P T H S Y A I N F N P Y I S				sag649_18rs21.pep
351	F G A T F T Q K A L M K A N E I L E T Q S S N A R K K L I F H V T D G V P T H S Y A I N F N P Y I S				sag649_nem316.pep
	Figure 23				

T S Y Q N Q E N S E L U N X I R D E S G I L E L E D E F E N G D D Y Q I V K G D G E S F K L F S D R K V Majority				
410	420	430	440	450
401 T S Y Q N Q F N S F L N K I P D R S G I L Q E D F I I N G D D Y Q I V K G D G E S F K L F S D R K V sag649_2603.pep				
401 T S Y Q N Q F N S F L N K I P D R S G I L Q E D F I I N G D D Y Q I V K G D G E S F K L F S D R K V sag649_coh1.pep				
401 T S Y Q N Q F N S F L N K I P D R S G I L Q E D F I I N G D D Y Q I V K G D G E S F K L F S D R K V sag649_cjb111.pep				
401 T S Y Q N Q F N S F L N K I P D R S G I L Q E D F I I N G D D Y Q I V K G D G E S F K L F S D R K V sag649_18rs21.pep				
401 T S Y Q N Q F N S F L N K I P D R S G I L Q E D F I I N G D D Y Q I V K G D G E S F K L F S D R K V sag649_nem316.pep				
P V T G G T T Q A A Y R V P Q N Q L S V M S N E G Y A I N S G Y I I Y L Y W R D Y N W V Y P F D P K T Majority				
460	470	480	490	500
451 P V T G G T T Q A A Y R V P Q N Q L S V M S N E G Y A I N S G Y I I Y L Y W R D Y N W V Y P F D P K T sag649_2603.pep				
451 P V T G G T T Q A A Y R V P Q N Q L S V M S N E G Y A I N S G Y I I Y L Y W R D Y N W V Y P F D P K T sag649_coh1.pep				
451 P V T G G T T Q A A Y R V P Q N Q L S V M S N E G Y A I N S G Y I I Y L Y W R D Y N W V Y P F D P K T sag649_cjb111.pep				
451 P V T G G T T Q A A Y R V P Q N Q L S V M S N E G Y A I N S G Y I I Y L Y W R D Y N W V Y P F D P K T sag649_18rs21.pep				
451 P V T G G T T Q A A Y R V P Q N Q L S V M S N E G Y A I N S G Y I I Y L Y W R D Y N W V Y P F D P K T sag649_nem316.pep				
K K V S A T K Q I K T H G E P T T L Y F N G N I R P K G Y D I F T V G I G V N G D P G A T P L E A E Majority				
510	520	530	540	550
501 K K V S A T K Q I K T H G E P T T L Y F N G N I R P K G Y D I F T V G I G V N G D P G A T P L E A E sag649_2603.pep				
501 K K V S A T K Q I K T H G E P T T L Y F N G N I R P K G Y D I F T V G I G V N G D P G A T P L E A E sag649_coh1.pep				
501 K K V S A T K Q I K T H G E P T T L Y F N G N I R P K G Y D I F T V G I G V N G D P G A T P L E A E sag649_cjb111.pep				
501 K K V S A T K Q I K T H G E P T T L Y F N G N I R P K G Y D I F T V G I G V N G D P G A T P L E A E sag649_18rs21.pep				
501 K K V S A T K Q I K T H G E P T T L Y F N G N I R P K G Y D I F T V G I G V N G D P G A T P L E A E sag649_nem316.pep				
K F M Q S I S S K T E N Y T N V D D T N K I Y D E L N K Y F K T I V E E K H S I V D G N V T D P M G Majority				
560	570	580	590	600
551 K F M Q S I S S K T E N Y T N V D D T N K I Y D E L N K Y F K T I V E E K H S I V D G N V T D P M G sag649_2603.pep				
551 K F M Q S I S S K T E N Y T N V D D T N K I Y D E L N K Y F K T I V E E K H S I V D G N V T D P M G sag649_coh1.pep				
551 K F M Q S I S S K T E N Y T N V D D T N K I Y D E L N K Y F K T I V E E K H S I V D G N V T D P M G sag649_cjb111.pep				
551 K F M Q S I S S K T E N Y T N V D D T N K I Y D E L N K Y F K T I V E E K H S I V D G N V T D P M G sag649_18rs21.pep				
551 K F M Q S I S S K T E N Y T N V D D T N K I Y D E L N K Y F K T I V E E K H S I V D G N V T D P M G sag649_nem316.pep				
E M I E F Q L K N G Q S F T H D D Y V L V G N D G S Q L K N G V A L G G P N S D G G I L K D V T V T Majority				
610	620	630	640	650
601 E M I E F Q L K N G Q S F T H D D Y V L V G N D G S Q L K N G V A L G G P N S D G G I L K D V T V T sag649_2603.pep				
601 E M I E F Q L K N G Q S F T H D D Y V L V G N D G S Q L K N G V A L G G P N S D G G I L K D V T V T sag649_coh1.pep				
601 E M I E F Q L K N G Q S F T H D D Y V L V G N D G S Q L K N G V A L G G P N S D G G I L K D V T V T sag649_cjb111.pep				
601 E M I E F Q L K N G Q S F T H D D Y V L V G N D G S Q L K N G V A L G G P N S D G G I L K D V T V T sag649_18rs21.pep				
601 E M I E F Q L K N G Q S F T H D D Y V L V G N D G S Q L K N G V A L G G P N S D G G I L K D V T V T sag649_nem316.pep				
Y D K T S Q T I K I N H L N L G S G Q K V V L T Y D V R L K D N Y I S N K F Y N T N N R T T L S P K Majority				
660	670	680	690	700
651 Y D K T S Q T I K I N H L N L G S G Q K V V L T Y D V R L K D N Y I S N K F Y N T N N R T T L S P K sag649_2603.pep				
651 Y D K T S Q T I K I N H L N L G S G Q K V V L T Y D V R L K D N Y I S N K F Y N T N N R T T L S P K sag649_coh1.pep				
651 Y D K T S Q T I K I N H L N L G S G Q K V V L T Y D V R L K D N Y I S N K F Y N T N N R T T L S P K sag649_cjb111.pep				
651 Y D K T S Q T I K I N H L N L G S G Q K V V L T Y D V R L K D N Y I S N K F Y N T N N R T T L S P K sag649_18rs21.pep				
651 Y D K T S Q T I K I N H L N L G S G Q K V V L T Y D V R L K D N Y I S N K F Y N T N N R T T L S P K sag649_nem316.pep				
S E K E P N T I R D F P I P K I R D V R E F P V L T I S N Q K K M G E V E F I K V N K D K H S E S L Majority				
710	720	730	740	750
701 S E K E P N T I R D F P I P K I R D V R E F P V L T I S N Q K K M G E V E F I K V N K D K H S E S L sag649_2603.pep				
701 S E K E P N T I R D F P I P K I R D V R E F P V L T I S N Q K K M G E V E F I K V N K D K H S E S L sag649_coh1.pep				
701 S E K E P N T I R D F P I P K I R D V R E F P V L T I S N Q K K M G E V E F I K V N K D K H S E S L sag649_cjb111.pep				
701 S E K E P N T I R D F P I P K I R D V R E F P V L T I S N Q K K M G E V E F I K V N K D K H S E S L sag649_18rs21.pep				
701 S E K E P N T I R D F P I P K I R D V R E F P V L T I S N Q K K M G E V E F I K V N K D K H S E S L sag649_nem316.pep				
L G A K F Q L Q I E K D F S G Y K Q F V P E G S D V T T K N D G K I Y E K A L Q D G N Y K L Y E I S Majority				
760	770	780	790	800
751 L G A K F Q L Q I E K D F S G Y K Q F V P E G S D V T T K N D G K I Y F K A L Q D G N Y K L Y E I S sag649_2603.pep				
751 L G A K F Q L Q I E K D F S G Y K Q F V P E G S D V T T K N D G K I Y F K A L Q D G N Y K L Y E I S sag649_coh1.pep				
751 L G A K F Q L Q I E K D F S G Y K Q F V P E G S D V T T K N D G K I Y F K A L Q D G N Y K L Y E I S sag649_cjb111.pep				
751 L G A K F Q L Q I E K D F S G Y K Q F V P E G S D V T T K N D G K I Y F K A L Q D G N Y K L Y E I S sag649_18rs21.pep				
751 L G A K F Q L Q I E K D F S G Y K Q F V P E G S D V T T K N D G K I Y F K A L Q D G N Y K L Y E I S sag649_nem316.pep				
751 L G A K F Q L Q I E K D F S G Y K Q F V P E G S D V T T K N D G K I Y F K A L Q D G N Y K L Y E I S sag649_cjb111.pep				

FIGURE 23A

Wednesday, May 20, 2004 \$0.00 1 W

	S P D G Y I E V K T K P V V T F T I Q N G E V T N L K A D P N A N K N Q I G Y L E G N G K H L I T N	Majority
810	820	830
801	S P D G Y I E V K T K P V V T F T I Q N G E V T N L K A D P N A N K N Q I G Y L E G N G K H L I T N	sag649_2603.pep
801	S P D G Y I E V K T K P V V T F T I Q N G E V T N L K A D P N A N K N Q I G Y L E G N G K H L I T N	sag649_coh1.pep
801	S P D G Y I E V K T K P V V T F T I Q N G E V T N L K A D P N A N K N Q I G Y L E G N G K H L I T N	sag649_cjb111.pep
801	S P D G Y I E V K T K P V V T F T I Q N G E V T N L K A D P N A N K N Q I G Y L E G N G K H L I T N	sag649_18rs21.pep
801	S P D G Y I E V K T K P V V T F T I Q N G E V T N L K A D P N A N K N Q I G Y L E G N G K H L I T N	sag649_nem316.pep
		840
		850
860	870	880
851	T P K R P P G V F P K T G G I G T I V Y I L V G S T F M I L T I C S F R R K Q L	sag649_2603.pep
851	T P K R P P G V F P K T G G I G T I V Y I L V G S T F M I L T I C S F R R K Q L	sag649_coh1.pep
851	T P K R P P G V F P K T G G I G T I V Y I L V G S T F M I L T I C S F R R K Q L	sag649_cjb111.pep
851	T P K R P P G V F P K T G G I G T I V Y I L V G S T F M I L T I C S F R R K Q L	sag649_18rs21.pep
851	T P K R P P G V F P K T G G I G T I V Y I L V G S T F M I L T I C S F R R K Q L	sag649_nem316.pep
		890
		Majority

Decoration #1: Shade (with solid black) residues that differ from the Consensus.

FIGURE 23B

137/487

Alignment Report WO 2006/078318 istal method with PAM250 residue weight table.
Thursday, July 29, 2004 7:08 PM.PCT/US2005/027239
Page 1

MR KY Q K F S P K I E T L I S I E F C L I S Q I E L N T N V L G E S T V P E N G A K G K L V V K K T D D Q Majority									
10	20	30	40	50					
1 M R K Y Q K F S K I L T L S L F C L S Q I P L N T N V L G E S T V P E N G A K G K L V V K K T D D Q sag1408_2603.pep									
1 M R K Y Q K F S K I L T L S L F C L S Q I P L N T N V L G E S T V P E N G A K G K L V V K K T D D Q sag1408_515.pep									
1 M R K Y Q K F S K I L T L S L F C L S Q I P L N T N V L G E S T V P E N G A K G K L V V K K T D D Q sag1408_18rs21.seq									
1 M R K Y Q K F S K I L T L S L F C L S Q I P L N T N V L G E S T V P E N G A K G K L V V K K T D D Q sag1408_cjb111.pep									
1 M R K Y Q K F S K I L T L S L F C L S Q I P L N T N V L G E S T V P E N G A K G K L V V K K T D D Q sag1408_h36b.pep									
1 M R K Y Q K F S K I L T L S L F C L S Q I P L N T N V L G E S T V P E N G A K G K L V V K K T D D Q sag1408_nem316.pep									
N K P L S K A T F V L K T T A H P E S K I E K V T A E L T G E A T F D N L I P G D Y T L S E E T A P Majority									
60	70	80	90	100					
51 N K P L S K A T F V L K T T A H P E S K I E K V T A E L T G E A T F D N L I P G D Y T L S E E T A P sag1408_2603.pep									
51 N K P L S K A T F V L K T T A H P E S K I E K V T A E L T G E A T F D N L I P G D Y T L S E E T A P sag1408_515.pep									
51 N K P L S K A T F V L K T T A H P E S K I E K V T A E L T G E A T F D N L I P G D Y T L S E E T A P sag1408_18rs21.seq									
51 N K P L S K A T F V L K T T A H P E S K I E K V T A E L T G E A T F D N L I P G D Y T L S E E T A P sag1408_cjb111.pep									
51 N K P L S K A T F V L K T T A H P E S K I E K V T A E L T G E A T F D N L I P G D Y T L S E E T A P sag1408_h36b.pep									
51 N K P L S K A T F V L K T T A H P E S K I E K V T A E L T G E A T F D N L I P G D Y T L S E E T A P sag1408_nem316.pep									
E G Y K E T N Q T W Q V K V E S N G K T T I Q N S G D K N S T I G Q N Q E E L D K Q Y P P T G I Y E Majority									
110	120	130	140	150					
101 E G Y K K T N Q T W Q V K V E S N G K T T I Q N S G D K N S T I G Q N Q E E L D K Q Y P P T G I Y E sag1408_2603.pep									
101 E G Y K K T N Q T W Q V K V E S N G K T T I Q N S G D K N S T I G Q N Q E E L D K Q Y P P T G I Y E sag1408_515.pep									
101 E G Y K K T N Q T W Q V K V E S N G K T T I Q N S G D K N S T I G Q N Q E E L D K Q Y P P T G I Y E sag1408_18rs21.seq									
101 E G Y K K T N Q T W Q V K V E S N G K T T I Q N S G D K N S T I G Q N Q E E L D K Q Y P P T G I Y E sag1408_cjb111.pep									
101 E G Y K K T N Q T W Q V K V E S N G K T T I Q N S G D K N S T I G Q N Q E E L D K Q Y P P T G I Y E sag1408_h36b.pep									
101 E G Y K K T N Q T W Q V K V E S N G K T T I Q N S G D K N S T I G Q N Q E E L D K Q Y P P T G I Y E sag1408_nem316.pep									
D T K E S Y K L E H V K G S V P N G K S E A K A V N P Y S S E G E H I R E I P E G T L S K R I S E V Majority									
160	170	180	190	200					
151 D T K E S Y K L E H V K G S V P N G K S E A K A V N P Y S S E G E H I R E I P E G T L S K R I S E V sag1408_2603.pep									
151 D T K E S Y K L E H V K G S V P N G K S E A K A V N P Y S S E G E H I R E I P E G T L S K R I S E V sag1408_515.pep									
151 D T K E S Y K L E H V K G S V P N G K S E A K A V N P Y S S E G E H I R E I P E G T L S K R I S E V sag1408_18rs21.seq									
151 D T K E S Y K L E H V K G S V P N G K S E A K A V N P Y S S E G E H I R E I P E G T L S K R I S E V sag1408_cjb111.pep									
151 D T K E S Y K L E H V K G S V P N G K S E A K A V N P Y S S E G E H I R E I P E G T L S K R I S E V sag1408_h36b.pep									
151 D T K E S Y K L E H V K G S V P N G K S E A K A V N P Y S S E G E H I R E I P E G T L S K R I S E V sag1408_nem316.pep									
G D L A H N K Y K I E L T V S G K T I V K P V D K Q K P L D V V F V L D N S N S M N N D G P N F Q R Majority									
210	220	230	240	250					
201 G D L A H N K Y K I E L T V S G K T I V K P V D K Q K P L D V V F V L D N S N S M N N D G P N F Q R sag1408_2603.pep									
201 G D L A H N K Y K I E L T V S G K T I V K P V D K Q K P L D V V F V L D N S N S M N N D G P N F Q R sag1408_515.pep									
201 G D L A H N K Y K I E L T V S G K T I V K P V D K Q K P L D V V F V L D N S N S M N N D G P N F Q R sag1408_18rs21.seq									
201 G D L A H N K Y K I E L T V S G K T I V K P V D K Q K P L D V V F V L D N S N S M N N D G P N F Q R sag1408_cjb111.pep									
201 G D L A H N K Y K I E L T V S G K T I V K P V D K Q K P L D V V F V L D N S N S M N N D G P N F Q R sag1408_h36b.pep									
201 G D L A H N K Y K I E L T V S G K T I V K P V D K Q K P L D V V F V L D N S N S M N N D G P N F Q R sag1408_nem316.pep									
H N K A K K A A E A L G T A V K D I L G A N S D N R V A L V T Y G S D I F D G R S V D V V K G F K E Majority									
260	270	280	290	300					
251 H N K A K K A A E A L G T A V K D I L G A N S D N R V A L V T Y G S D I F D G R S V D V V K G F K E sag1408_2603.pep									
251 H N K A K K A A E A L G T A V K D I L G A N S D N R V A L V T Y G S D I F D G R S V D V V K G F K E sag1408_515.pep									
251 H N K A K K A A E A L G T A V K D I L G A N S D N R V A L V T Y G S D I F D G R S V D V V K G F K E sag1408_18rs21.seq									
247 H N K A K K A A E A L G T A V K D I L G A N S D N R V A L V T Y G S D I F D G R S V D V V K G F K E sag1408_cjb111.pep									
251 H N K A K K A A E A L G T A V K D I L G A N S D N R V A L V T Y G S D I F D G R S V D V V K G F K E sag1408_h36b.pep									
251 H N K A K K A A E A L G T A V K D I L G A N S D N R V A L V T Y G S D I F D G R S V D V V K G F K E sag1408_nem316.pep									
D D K Y Y G L Q T K F T I Q T E N Y S H K Q L T N N A E E I I K R I P T E A P K A K W G S T T N G L Majority									
310	320	330	340	350					
301 D D K Y Y G L Q T K F T I Q T E N Y S H K Q L T N N A E E I I K R I P T E A P K A K W G S T T N G L sag1408_2603.pep									
301 D D K Y Y G L Q T K F T I Q T E N Y S H K Q L T N N A E E I I K R I P T E A P K A K W G S T T N G L sag1408_515.pep									
301 D D K Y Y G L Q T K F T I Q T E N Y S H K Q L T N N A E E I I K R I P T E A P K A K W G S T T N G L sag1408_18rs21.seq									
297 D D K Y Y G L Q T K F T I Q T E N Y S H K Q L T N N A E E I I K R I P T E A P K A K W G S T T N G L sag1408_cjb111.pep									
301 D D K Y Y G L Q T K F T I Q T E N Y S H K Q L T N N A E E I I K R I P T E A P K A K W G S T T N G L sag1408_h36b.pep									
301 D D K Y Y G L Q T K F T I Q T E N Y S H K Q L T N N A E E I I K R I P T E A P K A K W G S T T N G L sag1408_nem316.pep									
T P E Q Q K E Y Y L S K V G E T F T M K A F M E A D D I L S Q V N R N S Q K I I V H V T D G V P T R Majority									
360	370	380	390	400					
351 T P E Q Q K E Y Y L S K V G E T F T M K A F M E A D D I L S Q V N R N S Q K I I V H V T D G V P T R sag1408_2603.pep									
351 T P E Q Q K E Y Y L S K V G E T F T M K A F M E A D D I L S Q V N R N S Q K I I V H V T D G V P T R sag1408_515.pep									
351 T P E Q Q K E Y Y L S K V G E T F T M K A F M E A D D I L S Q V N R N S Q K I I V H V T D G V P T R sag1408_18rs21.seq									
351 T P E Q Q K E Y Y L S K V G E T F T M K A F M E A D D I L S Q V N R N S Q K I I V H V T D G V P T R sag1408_cjb111.pep									
346 T P E K K R E Y D L S K V G E T F T M K A F M E A D D I L S S T O R K S N K I I V H V T D G V P T R sag1408_h36b.pep									
351 T P E Q Q K E Y Y L S K V G E T F T M K A F M E A D D I L S Q V D R N S Q K I I V H V T D G V P T R sag1408_nem316.pep									

Figure 24

S Y A I N N F K L G A S Y E S Q F E Q M K K N G M E N K S N F L L T D K P E D I K G N G E S Y F L F Majority				
410	420	430	440	450
401 S Y A I N N F K L G A S Y E S Q F E Q M K K N G M E N K S N F L L T D K P E D I K G N G E S Y F L F sag1408_2603.pep				
401 S Y A I N N F K L G A S Y E S Q F E Q M K K N G M E N K S N F L L T D K P E D I K G N G E S Y F L F sag1408_515.pep				
401 S Y A I N N F K L G A S Y E S Q F E Q M K K N G M E N K S N F L L T D K P E D I K G N G E S Y F L F sag1408_18rs21.seq				
396 S Y A I N S F V K G S T Y A X Q F E R I K F E G Y L D K X N Y P I T D P E K I G N G E S Y F L F sag1408_cjb111.pep				
401 S Y A I N N F K L G A S Y E S Q F E Q M K K N G M E N K S N F L L T D K P E D I K G N G E S Y F L F sag1408_h36b.pep				
<u>P L D S Y Q T Q I I S G N L Q K L H Y L D L N L N Y P K G T I Y R N G P V K E H G T P T K L Y I N S Majority</u>				
460	470	480	490	500
451 P L D S Y Q T Q I I S G N L Q K L H Y L D L N L N Y P K G T I Y R N G P V K E H G T P T K L Y I N S sag1408_2603.pep				
451 P L D S Y Q T Q I I S G N L Q K L H Y L D L N L N Y P K G T I Y R N G P V K E H G T P T K L Y I N S sag1408_515.pep				
451 P L D S Y Q T Q I I S G N L Q K L H Y L D L N L N Y P K G T I Y R N G P V K E H G T P T K L Y I N S sag1408_18rs21.seq				
446 P L D S Y Q T Q I I S G N L Q K L H Y L D L N L N Y P K G T I Y R N G P V R E H G T P T K L Y I N S sag1408_cjb111.pep				
451 P L D S Y Q T Q I I S G N L Q K L H Y L D L N L N Y P K G T I Y R N G P V R E H G T P T K L Y I N S sag1408_h36b.pep				
<u>L K Q K N Y D I F N F G I D I S G F R Q V Y N E D Y K K N Q D G T F Q K L K E E A F E L S D G E I T Majority</u>				
510	520	530	540	550
501 L K Q K N Y D I F N F G I D I S G F R Q V Y N E D Y K K N Q D G T F Q K L K E E A F E L S D G E I T sag1408_2603.pep				
501 L K Q K N Y D I F N F G I D I S G F R Q V Y N E D Y K K N Q D G T F Q K L K E E A F E L S D G E I T sag1408_515.pep				
501 L K Q K N Y D I F N F G I D I S G F R Q V Y N E D Y K K N Q D G T F Q K L K E E A F E L S D G E I T sag1408_18rs21.seq				
496 L K Q K N Y D I F N F G I D I S G F R Q V Y N E D Y K K N Q D G T F Q K L K E E A F E L S D G E I T sag1408_cjb111.pep				
501 L K Q K N Y D I F N F G I D I S G F R Q V Y N E D Y K K N Q D G T F Q K L K E E A F E L S D G E I T sag1408_h36b.pep				
<u>E L M R S F S S K P E Y Y T P I V T S A D T S N N E I L S K I Q Q Q F E T I L T K E N S I V N G T I Majority</u>				
560	570	580	590	600
551 E L M R S F S S K P E Y Y T P I V T S A D T S N N E I L S K I Q Q Q F E T I L T K E N S I V N G T I sag1408_2603.pep				
551 E L M R S F S S K P E Y Y T P I V T S A D T S N N E I L S K I Q Q Q F E T I L T K E N S I V N G T I sag1408_515.pep				
551 E L M R S F S S K P E Y Y T P I V T S A D T S N N E I L S K I Q Q Q F E T I L T K E N S I V N G T I sag1408_18rs21.seq				
546 E L M R S F S S K P E Y Y T P I V T S A D T S N N E I L S K I Q Q Q F E T I L T K E N S I V N G T I sag1408_cjb111.pep				
551 E L M R S F S S K P E Y Y T P I V T S A D T S N N E I L S K I Q Q Q F E T I L T K E N S I V N G T I sag1408_h36b.pep				
<u>E D P M G D K I N L Q L G N G Q T L Q P S D Y T L Q G N D G S V M K D G I A T G G P N N D G G I L K Majority</u>				
610	620	630	640	650
601 E D P M G D K I N L Q L G N G Q T L Q P S D Y T L Q G N D G S V M K D G I A T G G P N N D G G I L K sag1408_2603.pep				
601 E D P M G D K I N L Q L G N G Q T L Q P S D Y T L Q G N D G S V M K D G I A T G G P N N D G G I L K sag1408_515.pep				
601 E D P M G D K I N L Q L G N G Q T L Q P S D Y T L Q G N D G S V M K D G I A T G G P N N D G G I L K sag1408_18rs21.seq				
596 E D P M G D K I N L Q L G N G Q T L Q P S D Y T L Q G N D G S V M K D G I A T G G P N N D G G I L K sag1408_cjb111.pep				
601 E D P M G D K I N L Q L G N G Q T L Q P S D Y T L Q G N D G S V M K D G I A T G G P N N D G G I L K sag1408_h36b.pep				
<u>G V K L E Y I G N K L Y V R G L N L G E G Q K V T L T Y D V K L D D S F I S N K F Y D T N G R T T L Majority</u>				
660	670	680	690	700
651 G V K L E Y I G N K L Y V R G L N L G E G Q K V T L T Y D V K L D D S F I S N K F Y D T N G R T T L sag1408_2603.pep				
651 G V K L E Y I G N K L Y V R G L N L G E G Q K V T L T Y D V K L D D S F I S N K F Y D T N G R T T L sag1408_515.pep				
651 G V K L E Y I G N K L Y V R G L N L G E G Q K V T L T Y D V K L D D S F I S N K F Y D T N G R T T L sag1408_18rs21.seq				
646 G V K L E Y I G N K L Y V R G L N L G E G Q K V T L T Y D V K L D D S F I S N K F Y D T N G R T T L sag1408_cjb111.pep				
651 G V K L E Y I G N K L Y V R G L N L G E G Q K V T L T Y D V K L D D S F I S N K F Y D T N G R T T L sag1408_h36b.pep				
<u>N P K S E D P N T L R D F P I P K I R D V R E Y P T I T I K N E K K L G E F E F I K V D K D N N K L Majority</u>				
710	720	730	740	750
701 N P K S E D P N T L R D F P I P K I R D V R E Y P T I T I K N E K K L G E I E F I K V D K D N N K L sag1408_2603.pep				
701 N P K S E D P N T L R D F P I P K I R D V R E Y P T I T I K N E K K L G E I E F I K V D K D N N K L sag1408_515.pep				
701 N P K S E D P N T L R D F P I P K I R D V R E Y P T I T I K N E K K L G E I E F I K V D K D N N K L sag1408_18rs21.seq				
696 N P K S E D P N T L R D F P I P K I R D V R E Y P T I T I K N E K K L G E I E F I K V D K D N N K L sag1408_cjb111.pep				
701 N P K S E D P N T L R D F P I P K I R D V R E Y P T I T I K N E K K L G E I E F I K V D K D N N K L sag1408_h36b.pep				
<u>L L K G A T F E L Q E F N E D Y K L Y L P I K N N N S K V V T G E N G K I S Y K D L K D G K Y Q L I Majority</u>				
760	770	780	790	800
751 L L K G A T F E L Q E F N E D Y K L Y L P I K N N N S K V V T G E N G K I S Y K D L K D G K Y Q L I sag1408_2603.pep				
751 L L K G A T F E L Q E F N E D Y K L Y L P I K N N N S K V V T G E N G K I S Y K D L K D G K Y Q L I sag1408_515.pep				
751 L L K G A T F E L Q E F N E D Y K L Y L P I K N N N S K V V T G E N G K I S Y K D L K D G K Y Q L I sag1408_18rs21.seq				
746 L L K G A T F E L Q E F N E D Y K L Y L P I K N N N S K V V T G E N G K I S Y K D L K D G K Y Q L I sag1408_cjb111.pep				
751 L L K G A T F E L Q E F N E D Y K L Y L P I K N N N S K V V T G E N G K I S Y K D L K D G K Y Q L I sag1408_h36b.pep				
<u>L L K G A T F E L Q E F N E D Y K L Y L P I K N N N S K V V T G E N G K I S Y K D L K D G K Y Q L I sag1408_nem316.pep</u>				

FIGURE 24A

139/487

Alignment Report WO 2006/078318 istal method with PAM250 residue weight table.
Thursday, July 29, 2004 7:08 PM

PCT/US2005/027239
Page 3

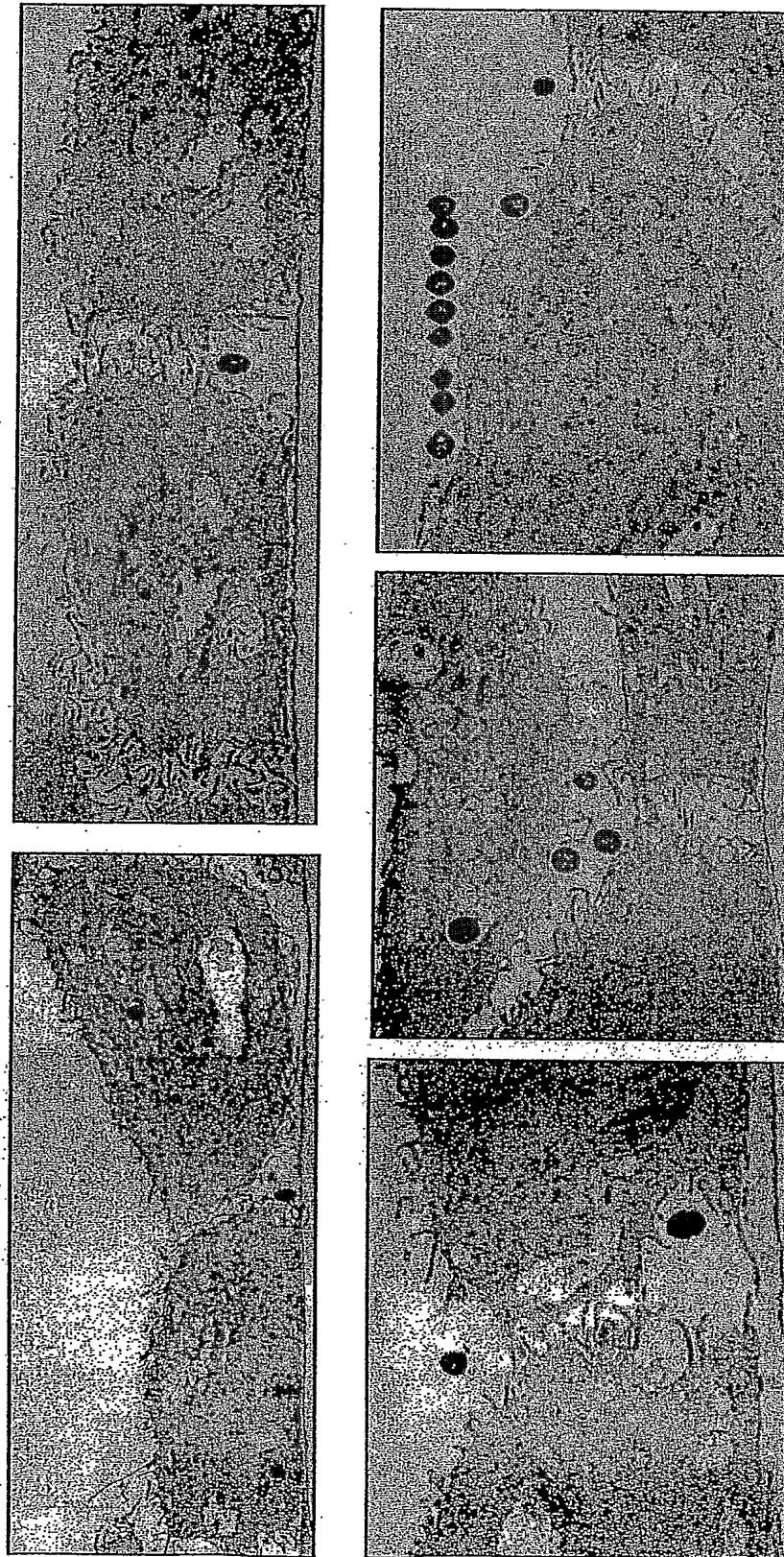
Page 3

E A V S P E D Y Q K I T N K P I L T F E V V K G S I N N I I A V N K Q I S E Y H E E G D K H L I T N Majority									
	810	820	830	840	850				
801	E A V S P E D Y Q K I T N K P I L T F E V V K G S I N N I I A V N K Q I S E Y H E E G D K H L I T N sagi408_2603.pep								
801	E A V S P E D Y Q K I T N K P I L T F E V V K G S I N N I I A V N K Q I S E Y H E E G D K H L I T N sagi408_515.pep								
801	E A V S P E D Y Q K I T N K P I L T F E V V K G S I N N I I A V N K Q I S E Y H E E G D K H L I T N sagi408_18rs21.seq								
801	E A V S P R D Y Q K I T N K P I L T F E V V K G S I Q N I I A V N K Q I S E Y H E E G D K H L I T N sagi408_cjb111.pep								
796	E A V S P R D Y Q K I T N K P I L T F E V V K G S I Q N I I A V N K Q I S E Y H E E G D K H L I T N sagi408_h36b.pep								
801	E A V S P R D Y Q K I T N K P I L T F E V V K G S I Q N I I A V N K Q I S E Y H E E G D K H L I T N sagi408_nem316.pep								
T H I P P K G I I P M T G G K G I L S F I L I G G A M M S I A G G I Y I W K R Y K K S S D M S I E K Majority									
	860	870	880	890	900				
851	T H I P P K G I I P M T G G K G I L S F I L I G G A M M S I A G G I Y I W K R Y K K S S D M S I E K sagi408_2603.pep								
851	T H I P P K G I I P M T G G K G I L S F I L I G G A M M S I A G G I Y I W K R Y K K S S D M S I E K sagi408_515.pep								
851	T H I P P K G I I P M T G G K G I L S F I L I G G A M M S I A G G I Y I W K R Y K K S S D M S I E K sagi408_18rs21.seq								
851	T H I P P K G I I P M T G G K G I L S F I L I G G S M M S I A G G I Y I W K R Y K K S S D I S R E K sagi408_cjb111.pep								
846	T H I P P K G I I P M T G G K G I L S F I L I G G A M M S I A G G I Y I W K R Y K K S S D A S I E K sagi408_h36b.pep								
851	T H I P P K G I I P M T G G K G I L S F I L I G G S M M S I A G G I Y I W K R Y K K S S D I S R E K sagi408_nem316.pep								
D									
									Majority
901	D								sagi408_2603.pep
901	D								sagi408_515.pep
901	D								sagi408_18rs21.seq
901	D								sagi408_cjb111.pep
896	D								sagi408_h36b.pep
901	D								sagi408_nem316.pep

Decoration 'Decoration #1': Shade (with solid black) residues that differ from the Consensus

FIGURE 24B

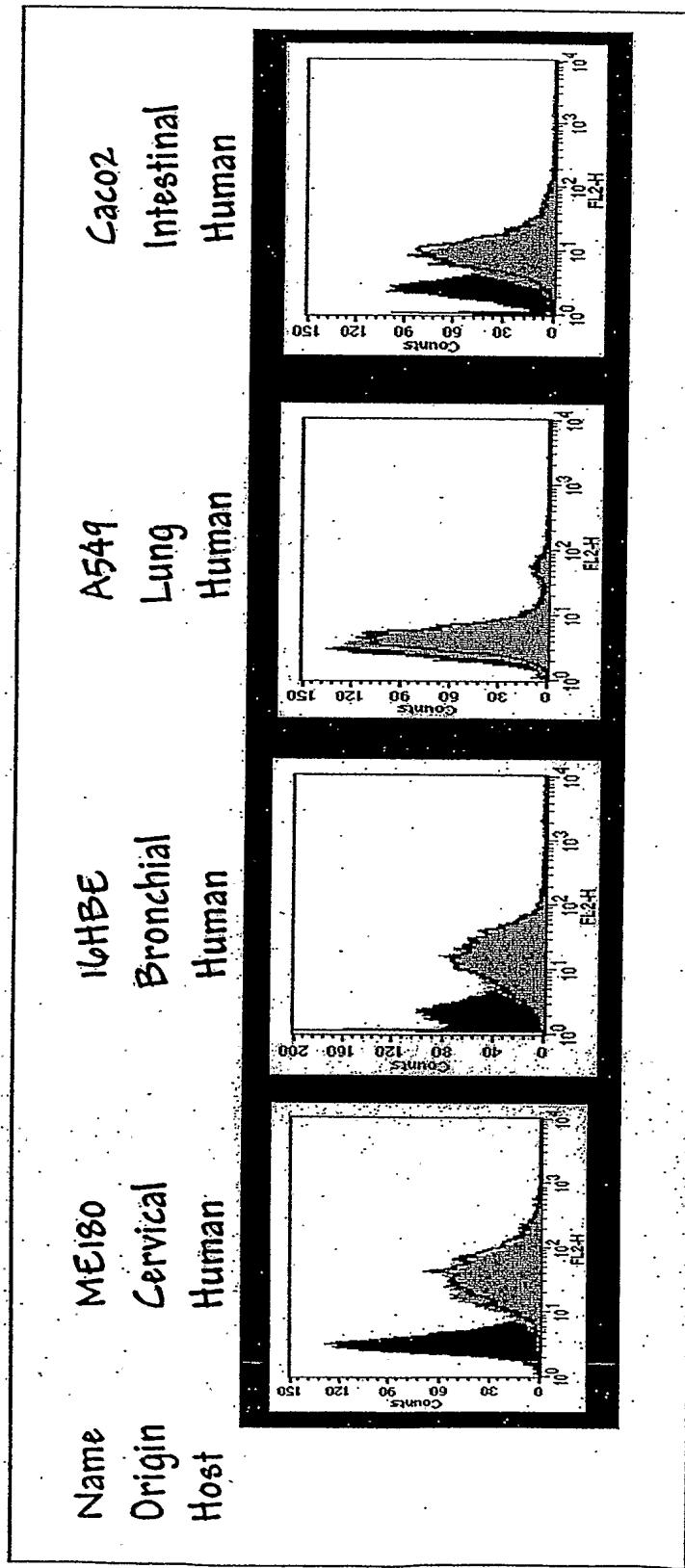
Figure 25. GBS closely associate with tight junctions and cross the monolayer by a paracellular route



Transmission Electron Microscopy images of GBS infection
of ME180 cervical epithelial cells.

Figure 26: GBS infection of ME180 cells



Figure 27

PCT/US605/27239 143/487

Figure 28

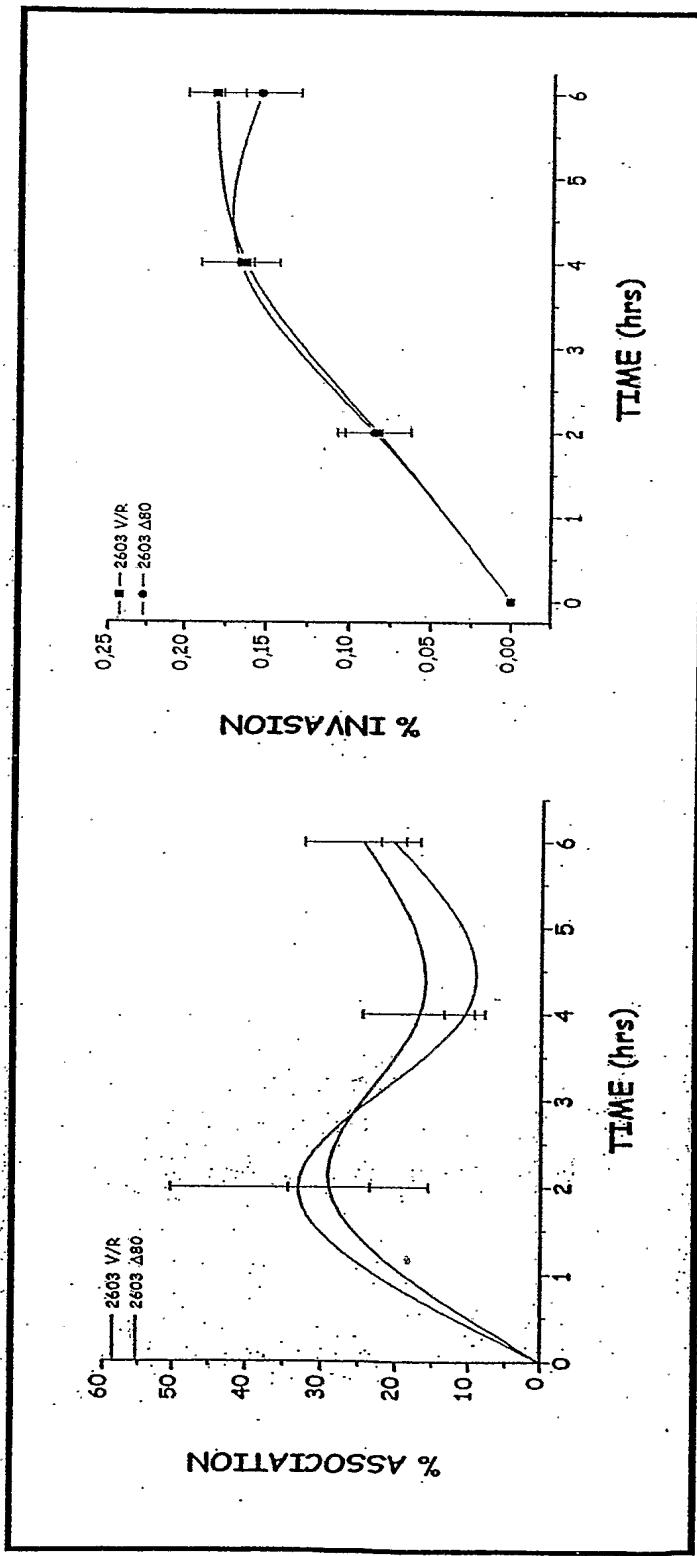


Figure 29

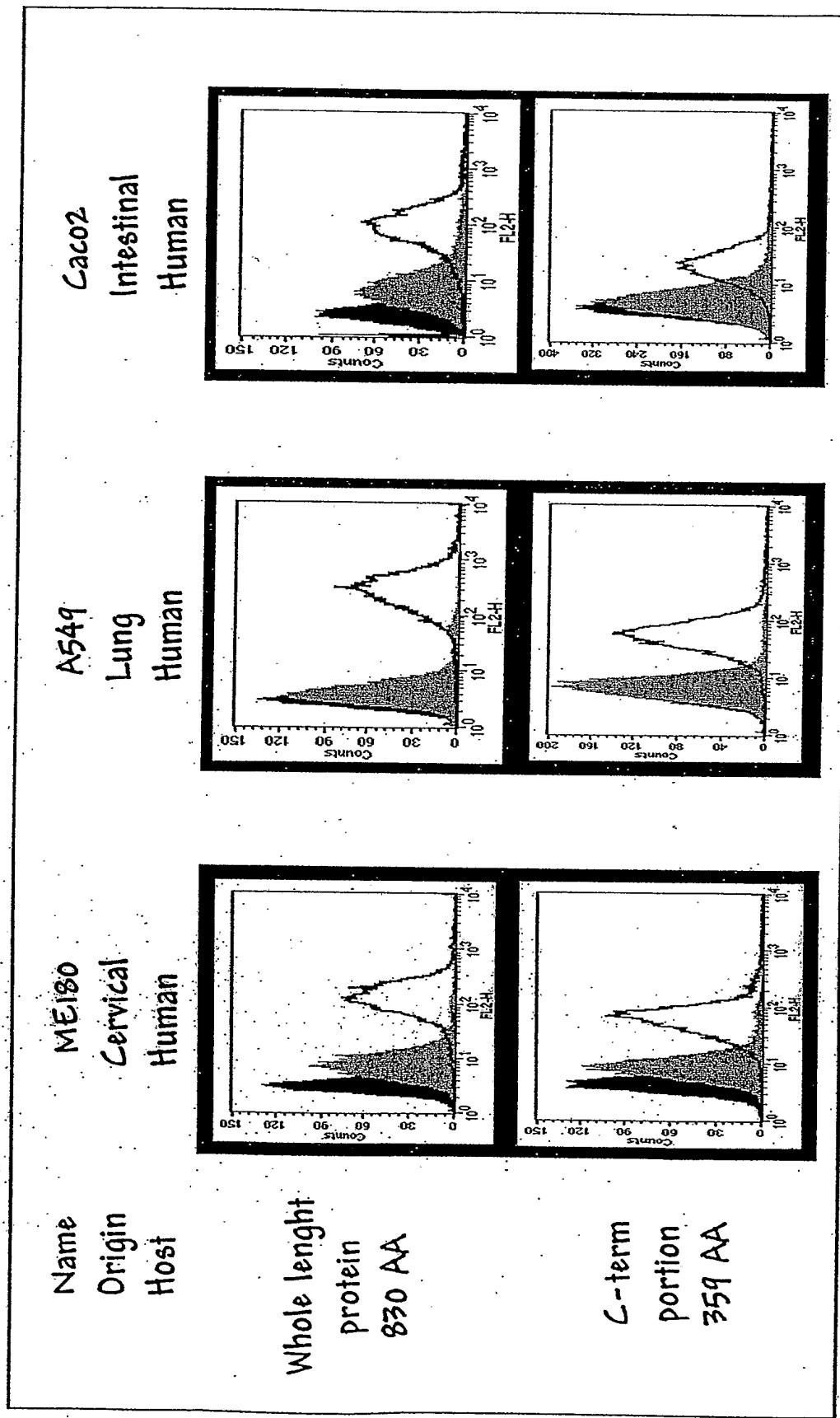


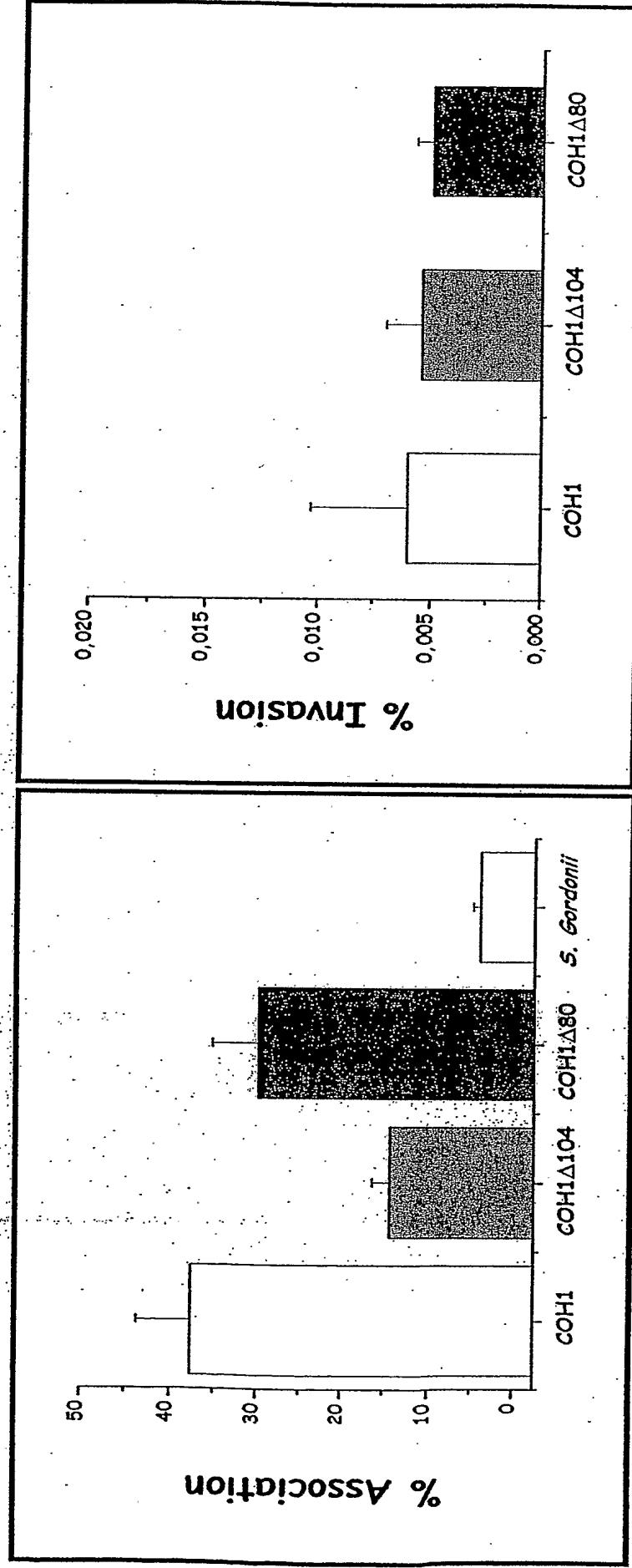
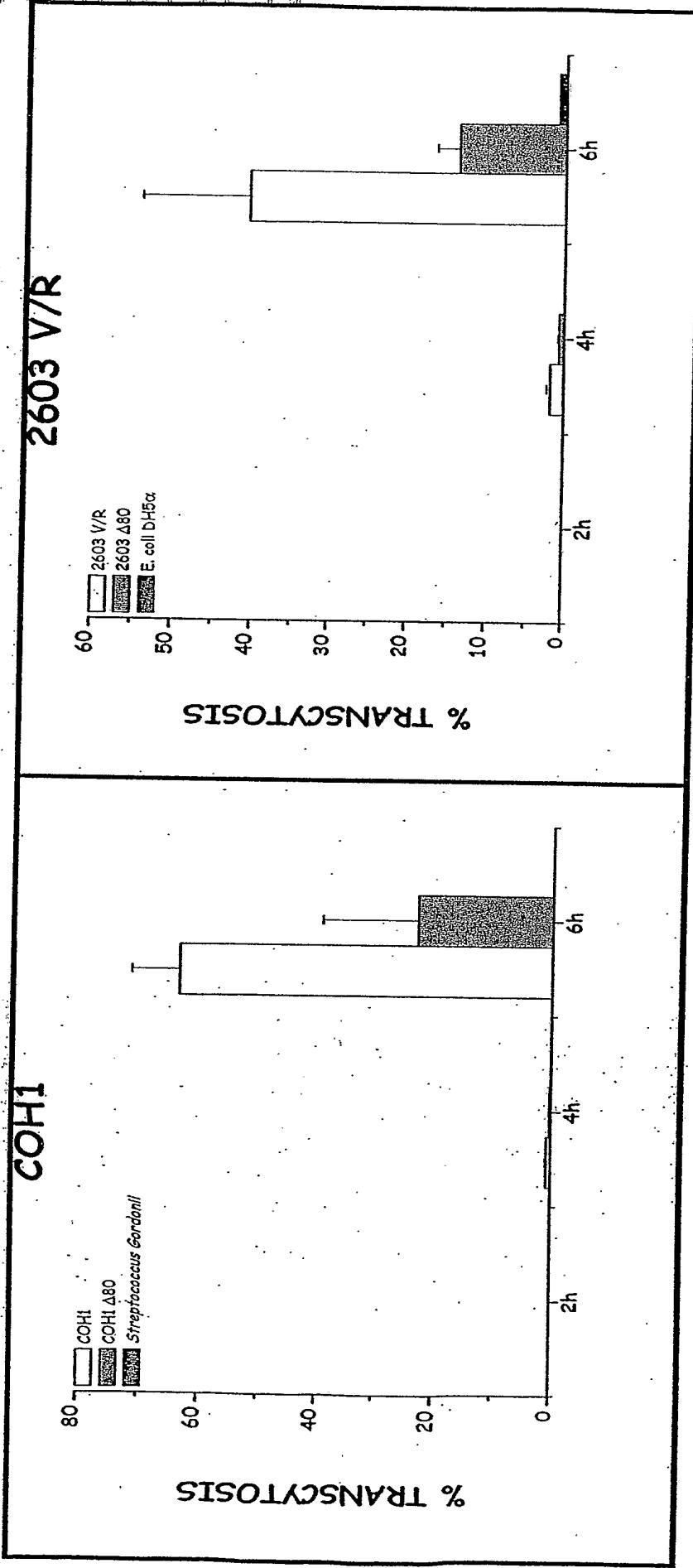
Figure 30

Figure 31

PCT/US2005/27239 147/487

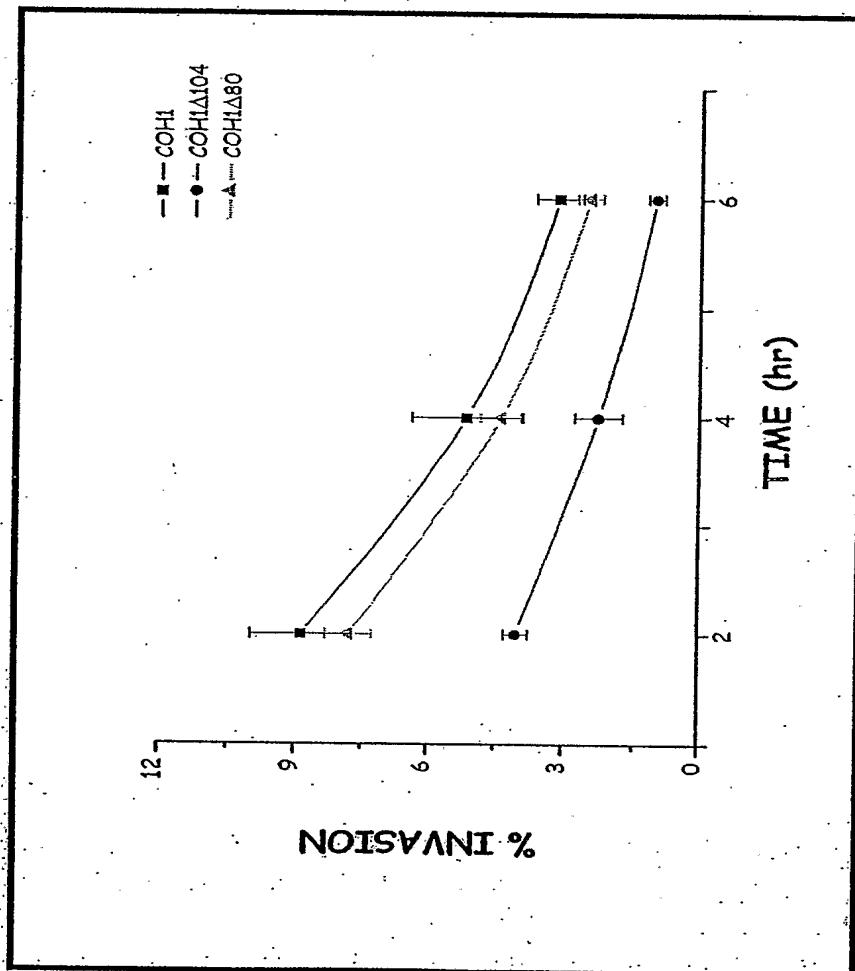
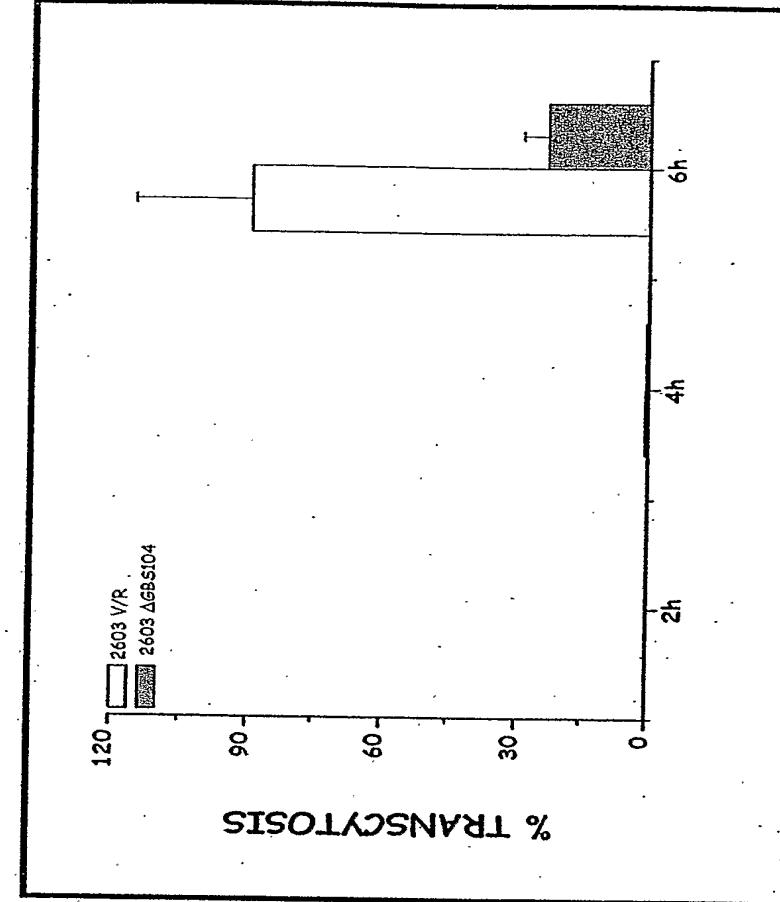
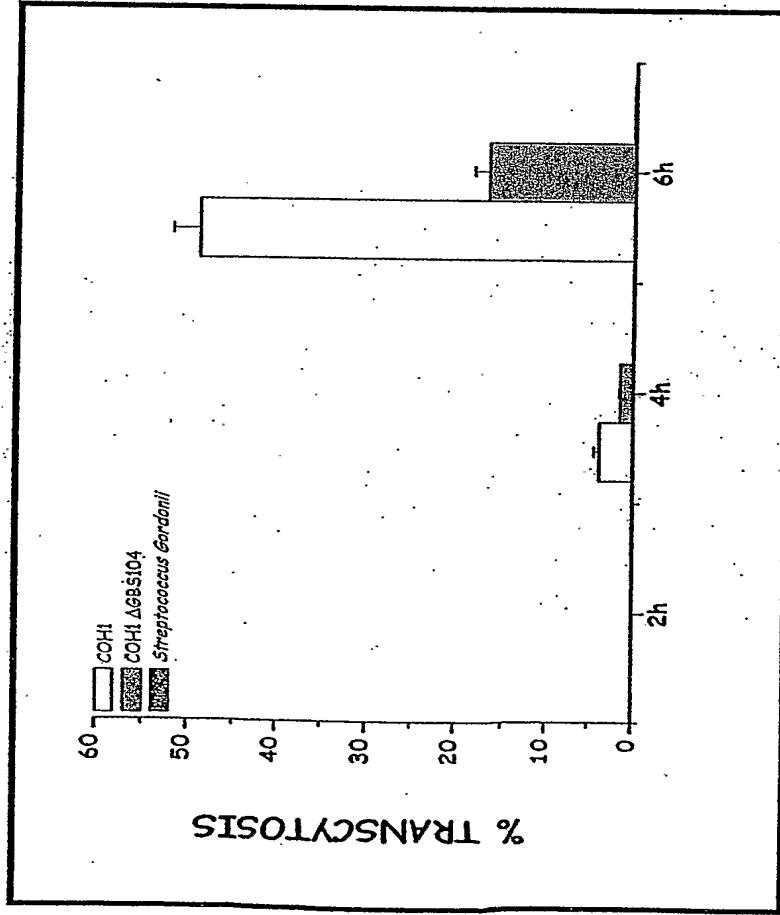
Figure 32

Figure 33

COH1

2603 V/R

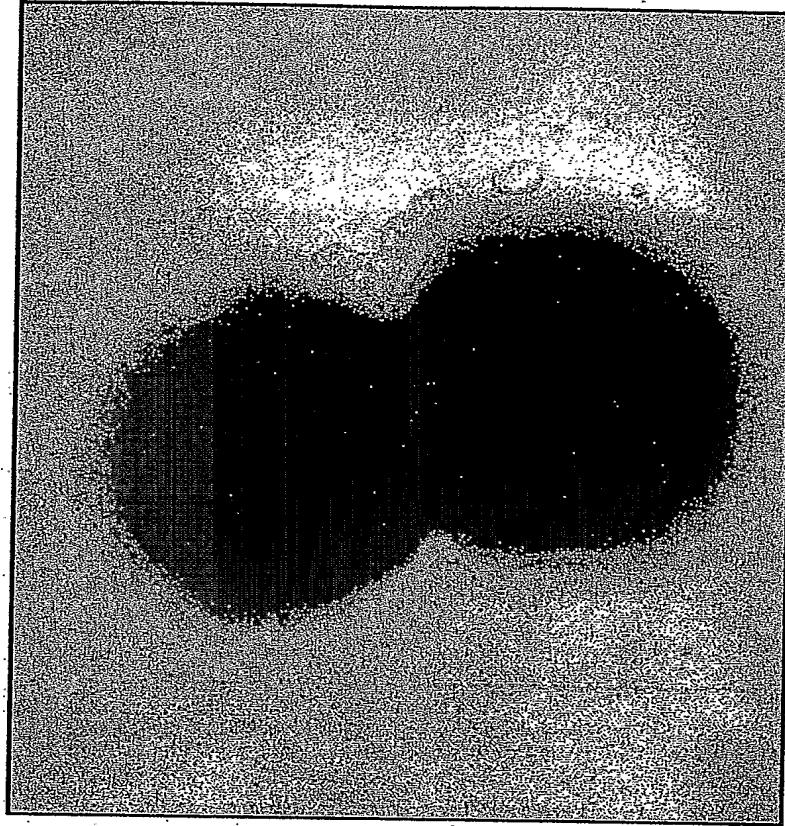
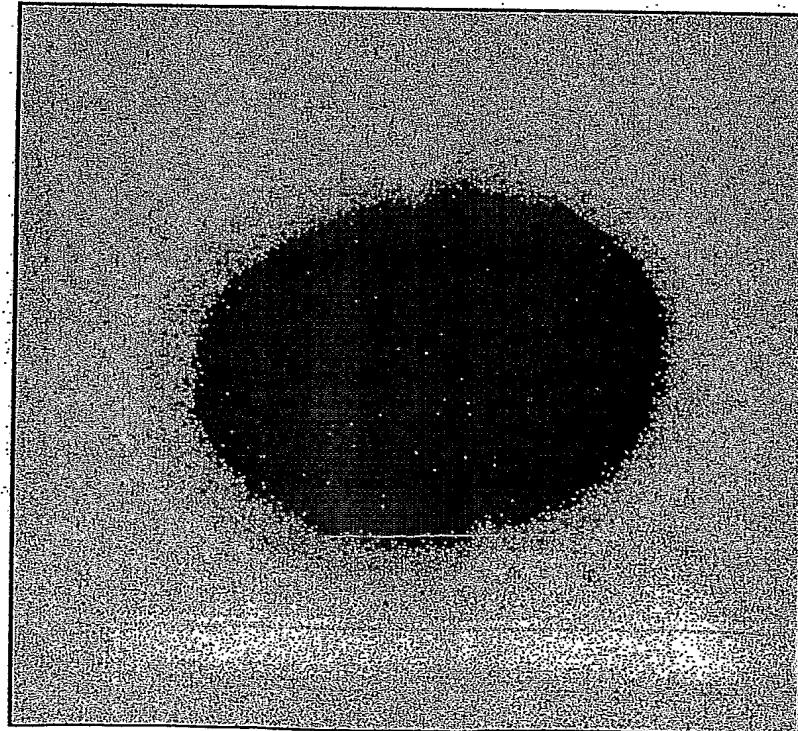


PCT/US05/27239 149/487

GBS STRAIN COH1 over GBS80

Figure 34

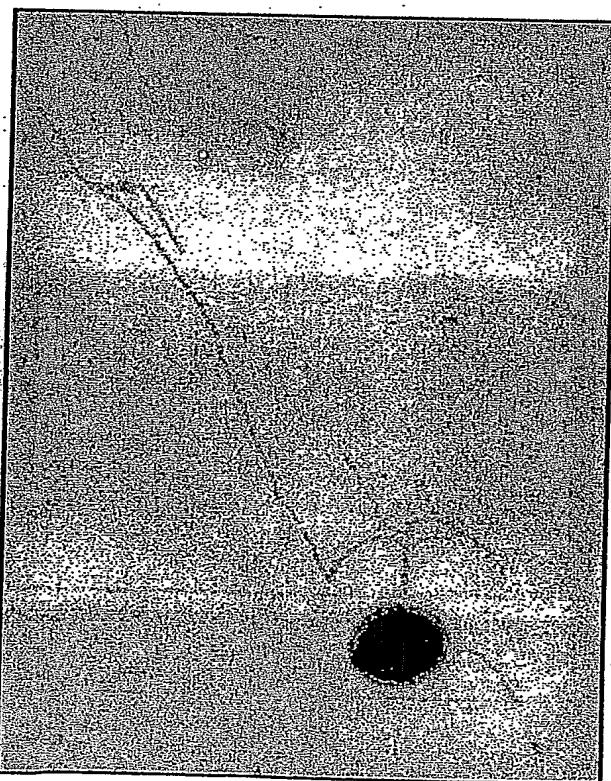
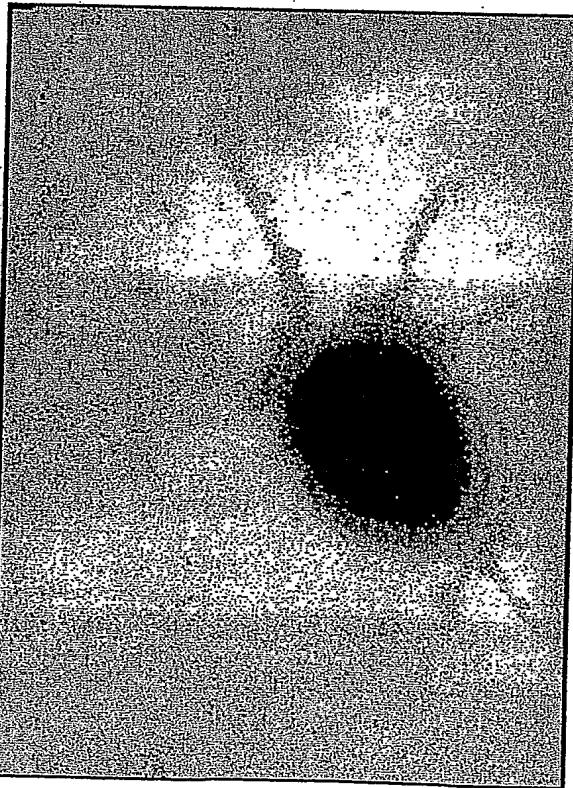
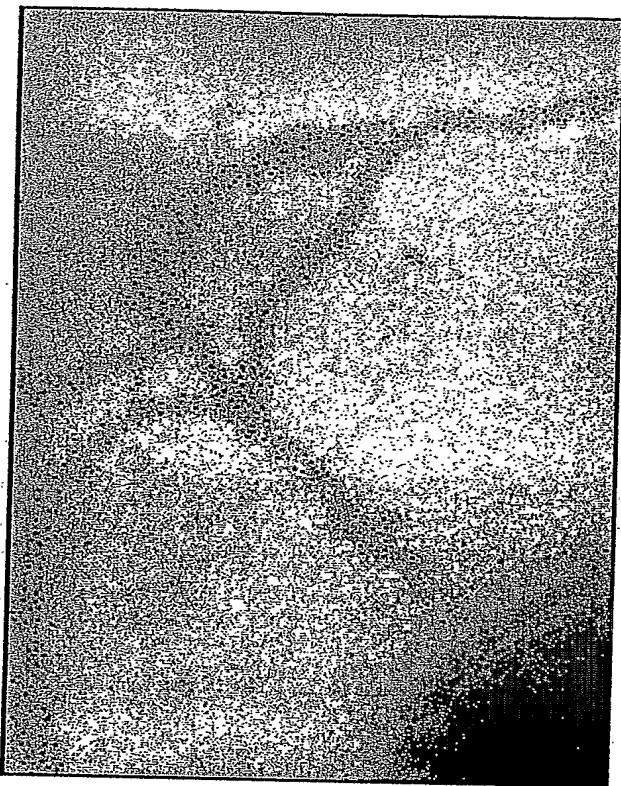
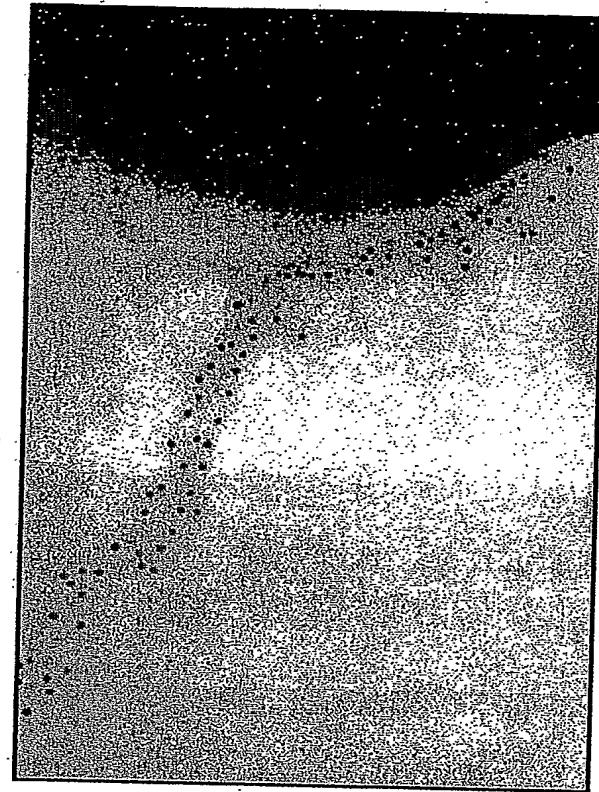
Negative staining EM



GBS STRAIN COH1 over GBS80

Figure 35

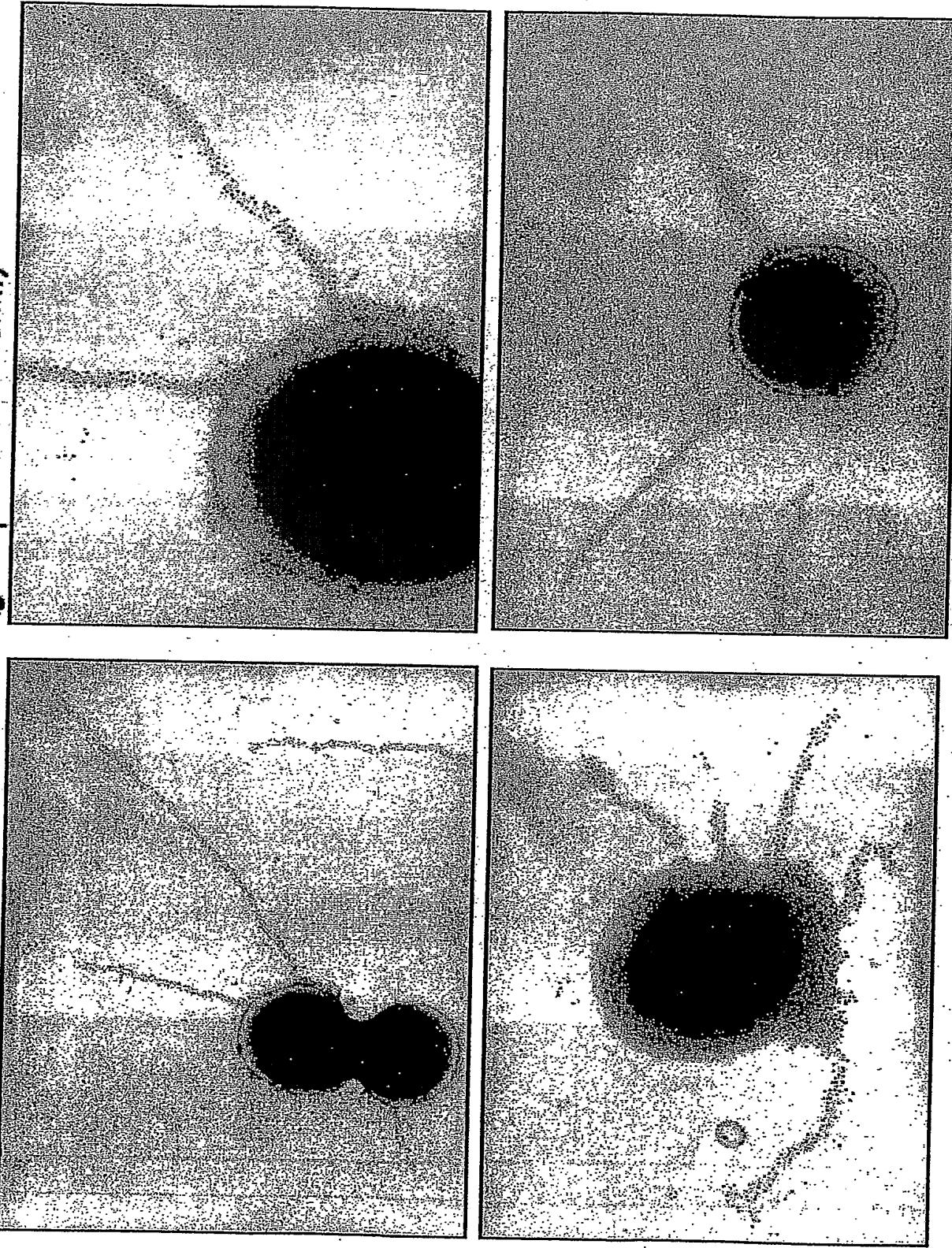
TEM anti-GBS80 (gold particles 10nm)



GBS STRAIN COH1 over GBS80

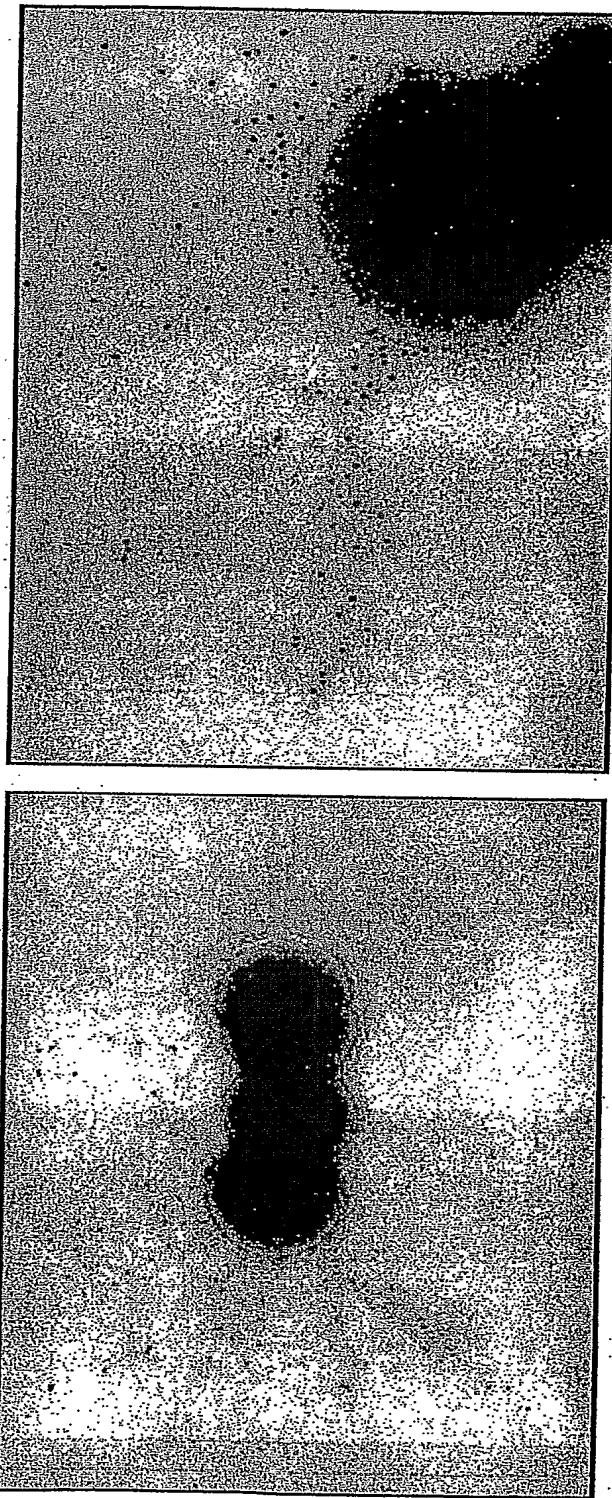
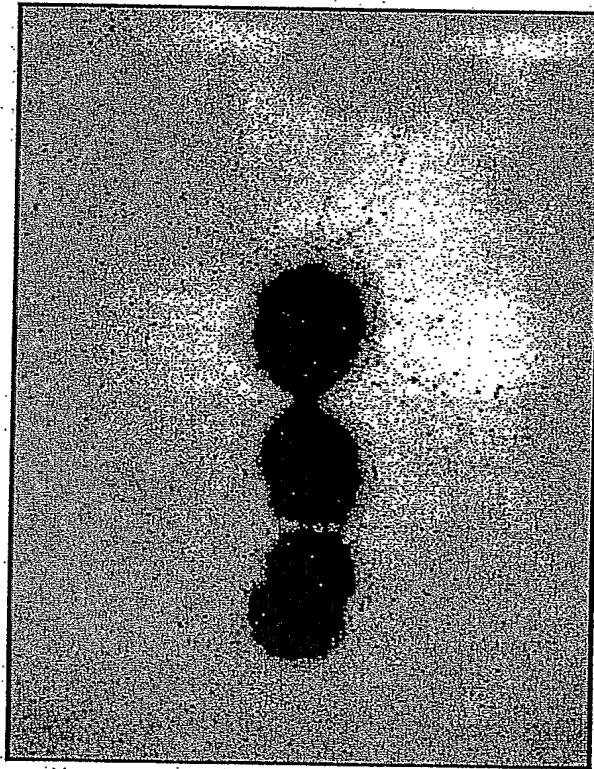
Figure 36

TEM anti-GBS80 (gold particles 10nm)



GBS STRAIN COH1 over GBS80
TEM anti-GBS80 (gold particles 20nm)

Figure 37



GBS STRAIN COH1 over GBS80

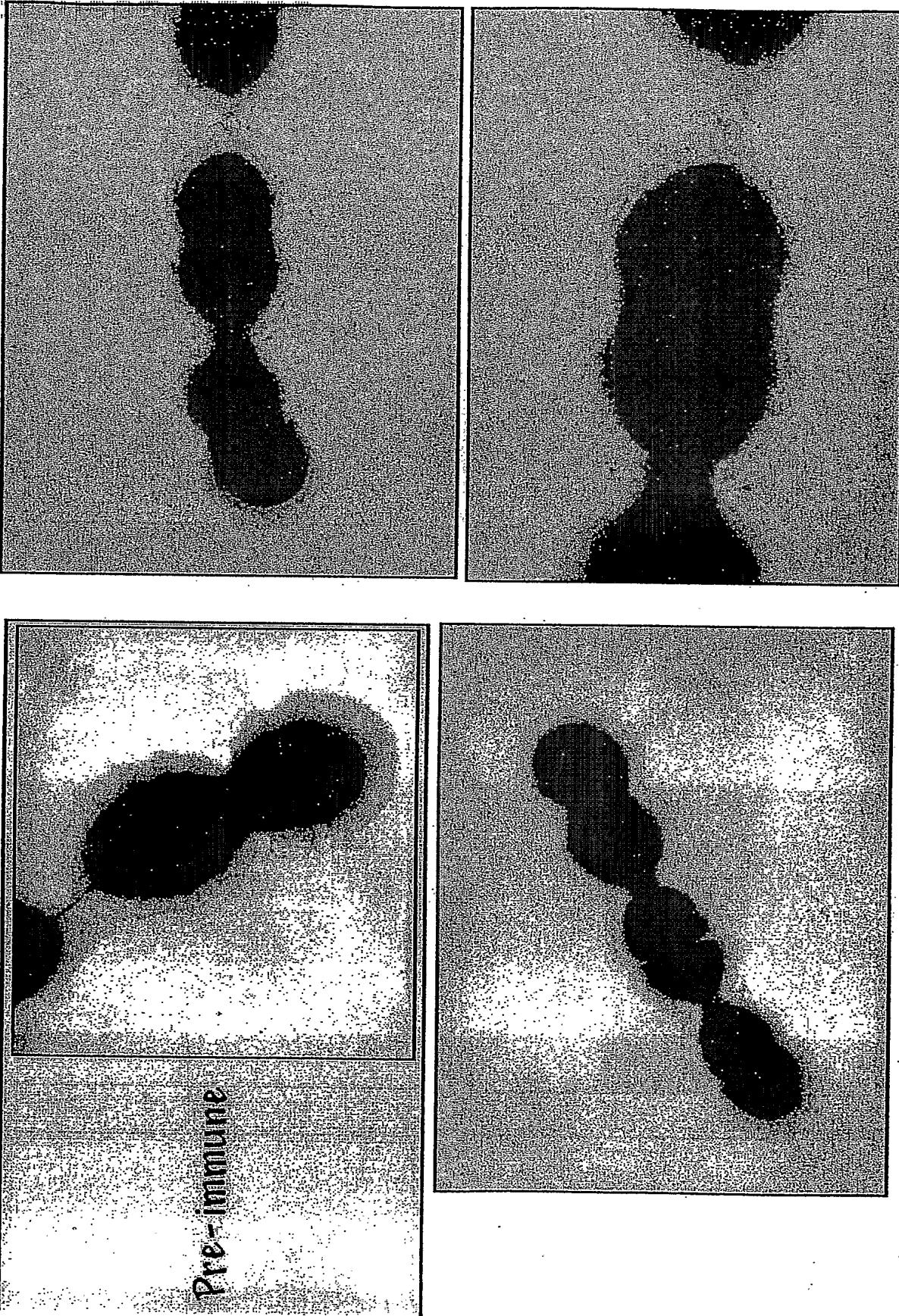
Figure 38

TEM anti-GBS104 (Gold particles 10nm)

WO 2006/078318

153/487

PCT/US2005/027239



GBS STRAIN COH1 over GBS80

Figure 39 IEM anti-GBS80 (gold particles 20nm) anti-GBS104 (gold particles 10nm)

WO 2006/078318

PCT/US05/27239

154/487

PCT/US2005/027239

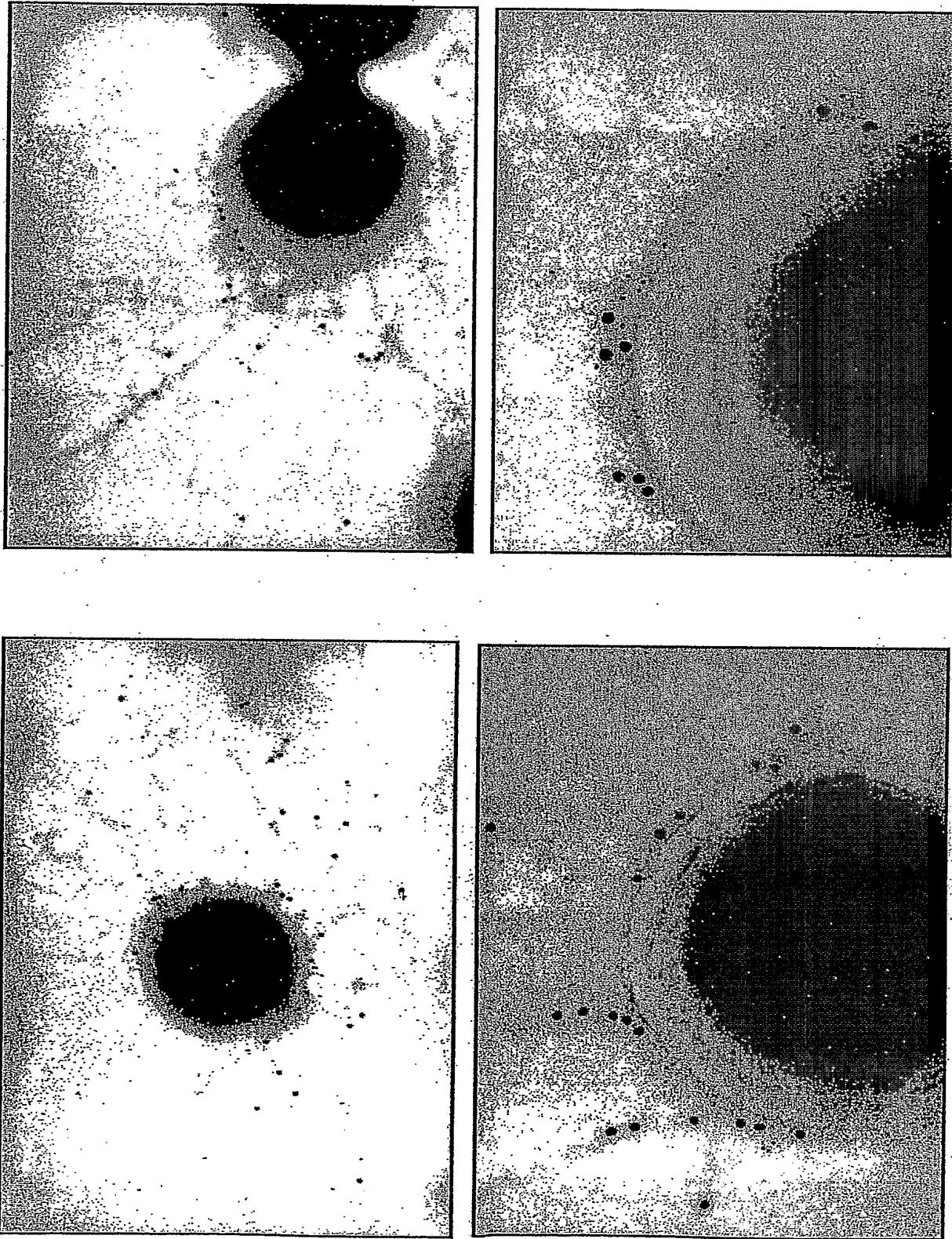
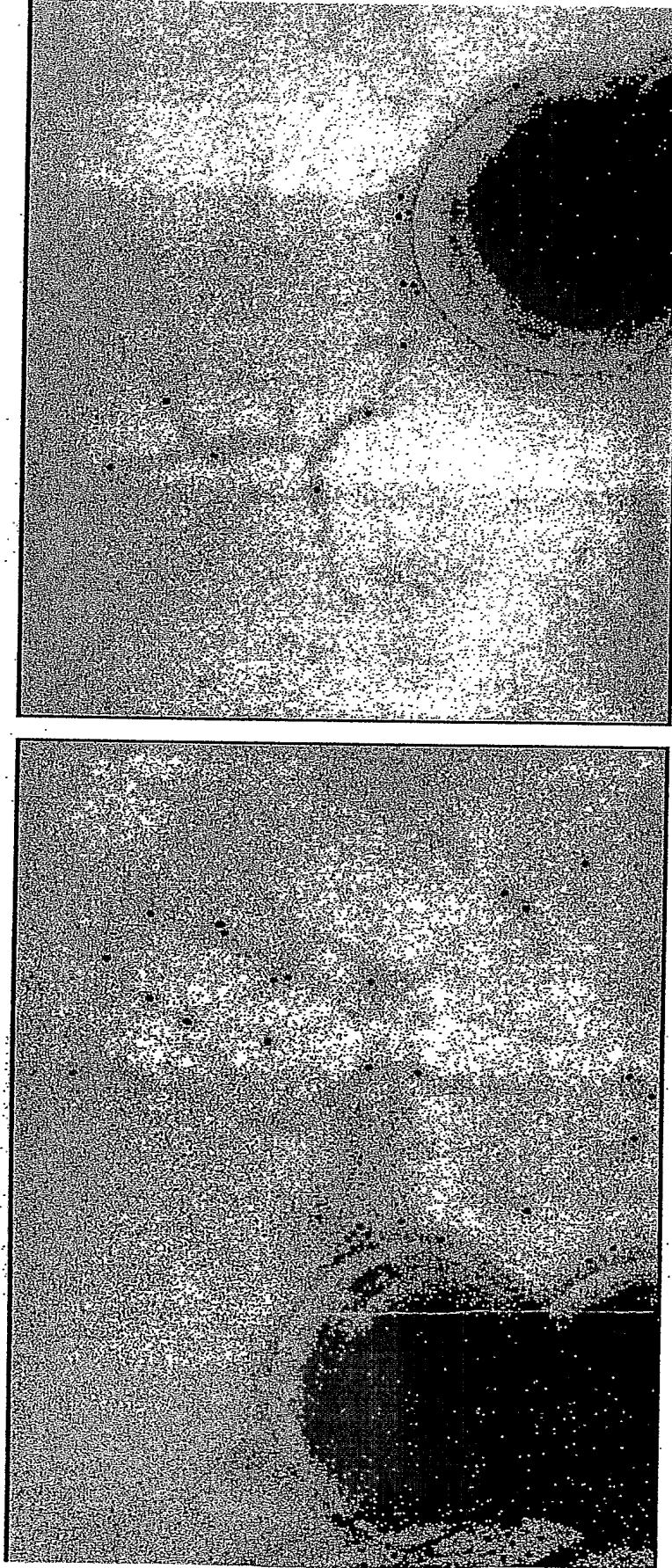


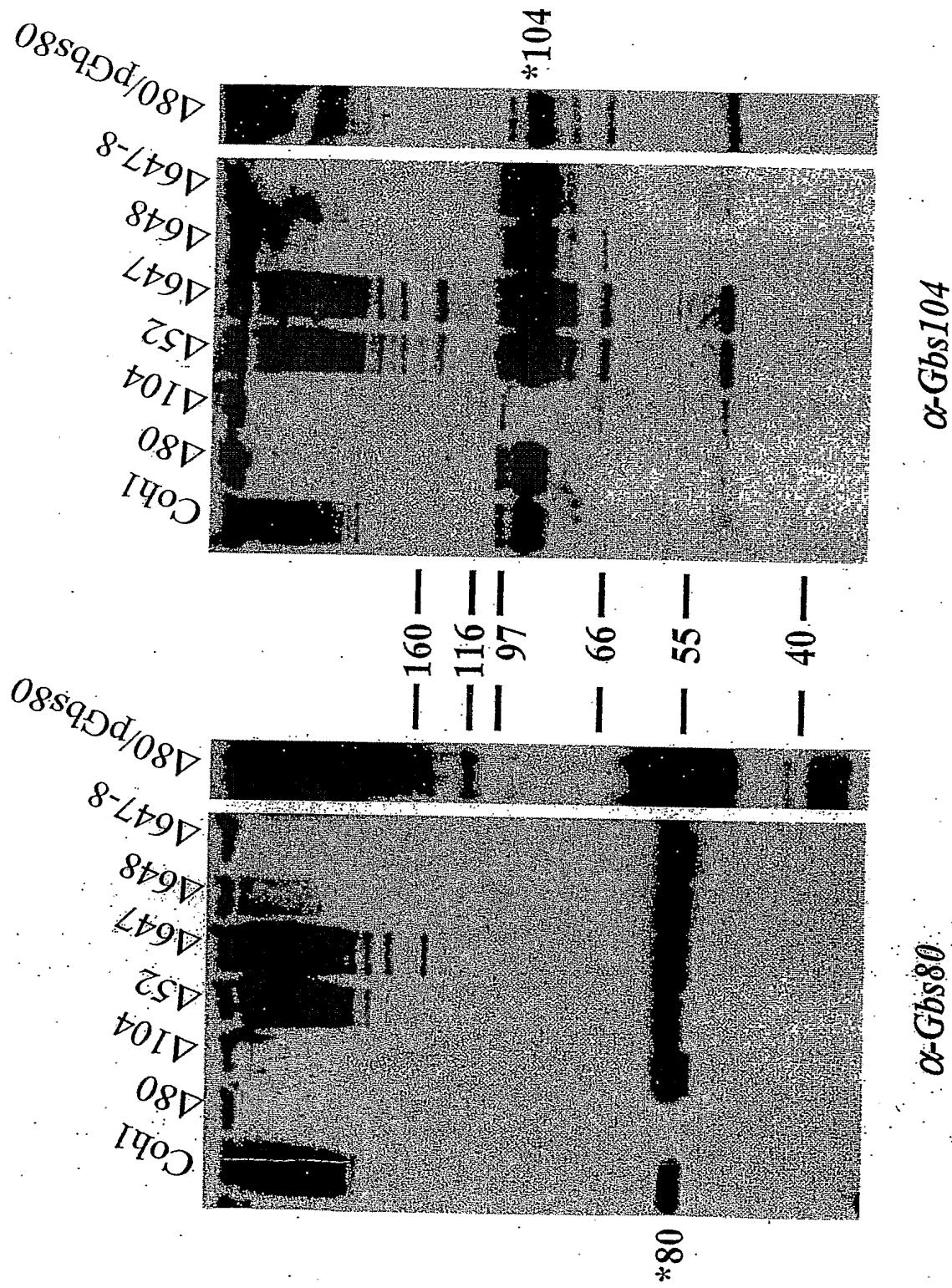
Figure 40 ***GBS STRAIN COH1 over GBS80***

TEM anti-GBS80 (gold particles 20nm) anti-GBS104 (gold particles 10nm)



156/487

Figure 41: GBS 80 is necessary for polymer formation, GBS104 and sortase SAG0648 are necessary for efficient assembly



PCT/US05/23239 157/487

**Figure 42: Gbs67 is part of a second pilus;
Gbs80 is polymerized in strain 515
(515 lacks sortase 647-8, but has AI-2 sortases)**

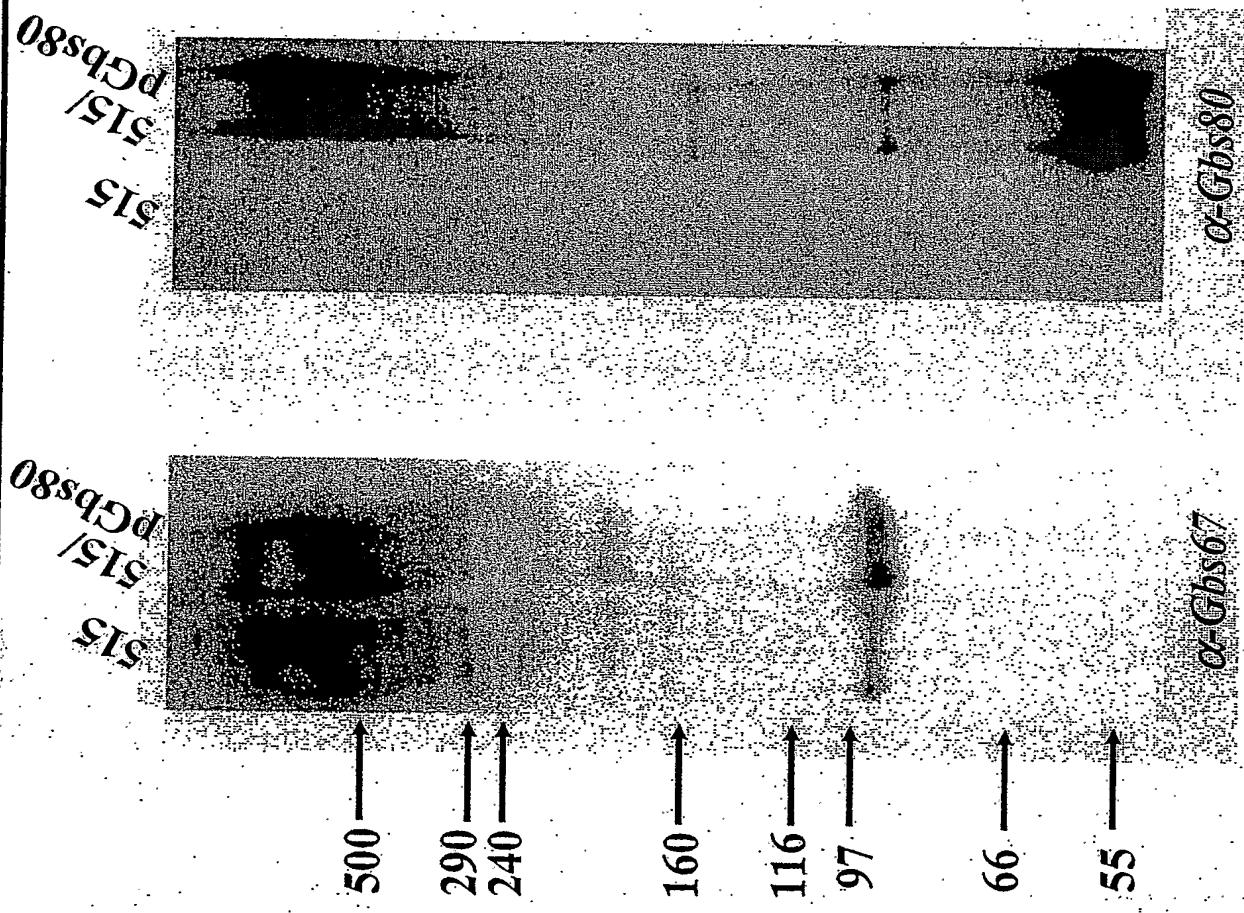
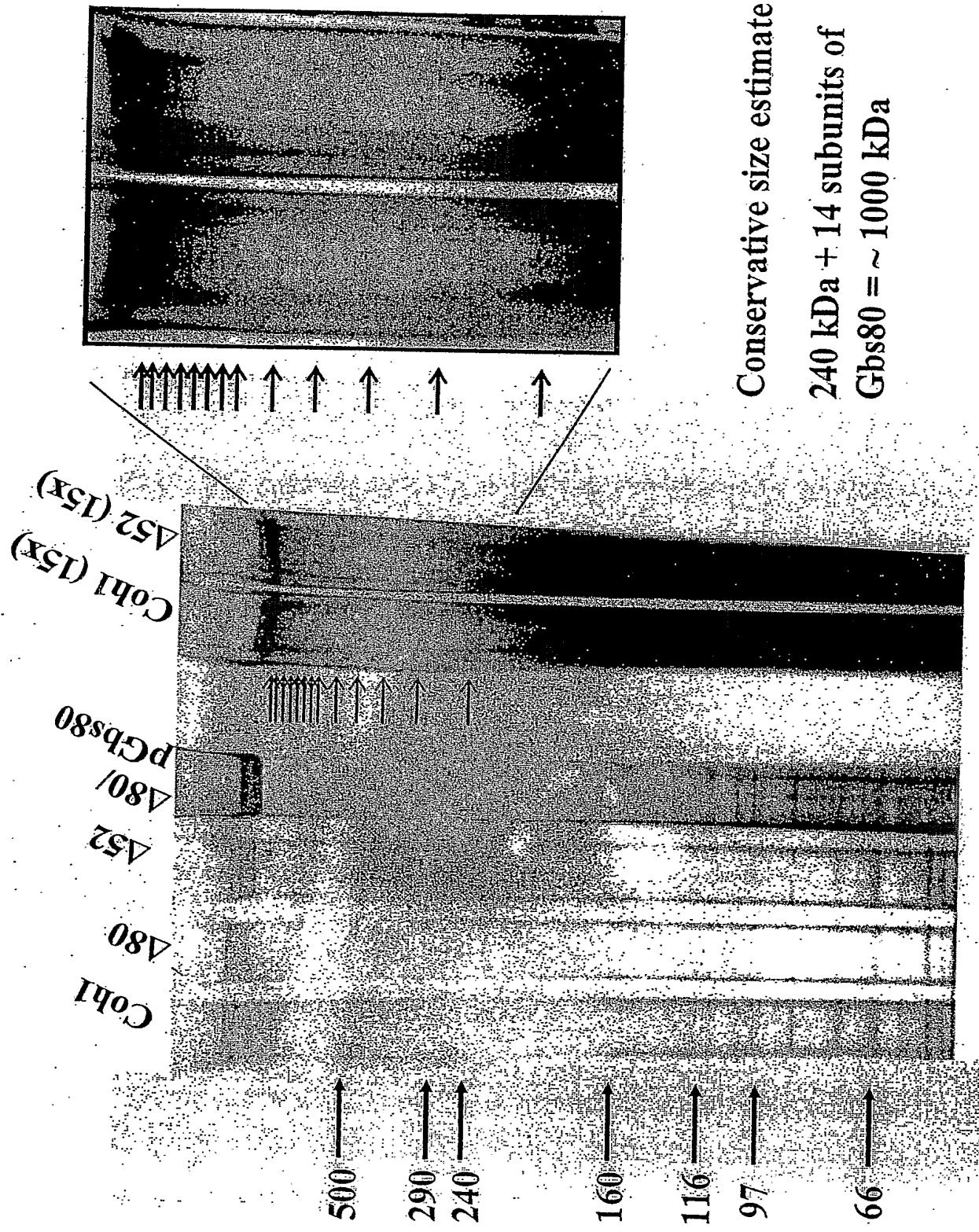


Figure 43: Two macro-molecules are visible in Coh1 at >10000 kDa, one is the Gbs80 pilin



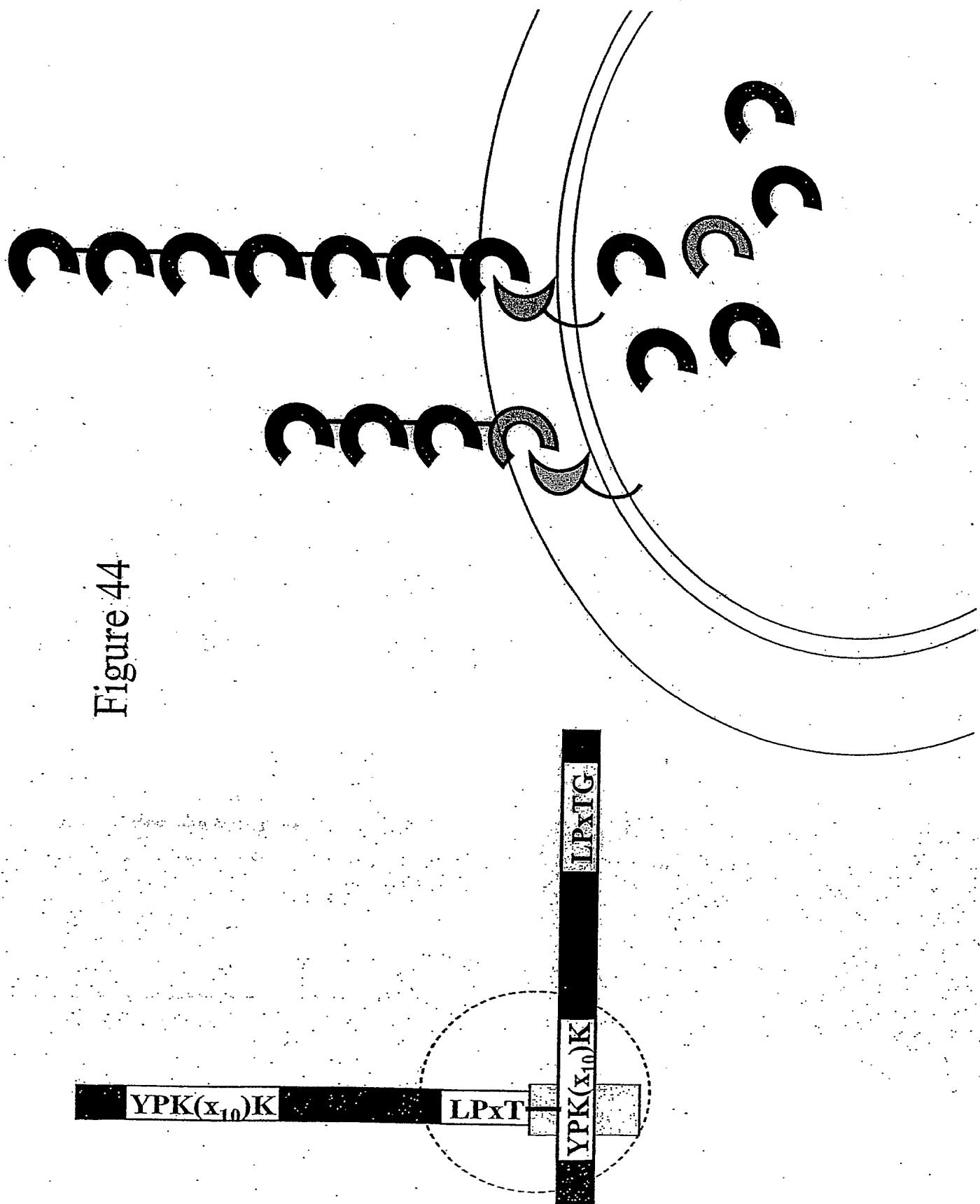


Figure 45: Gbs52 is a minor component of the GBs pilus

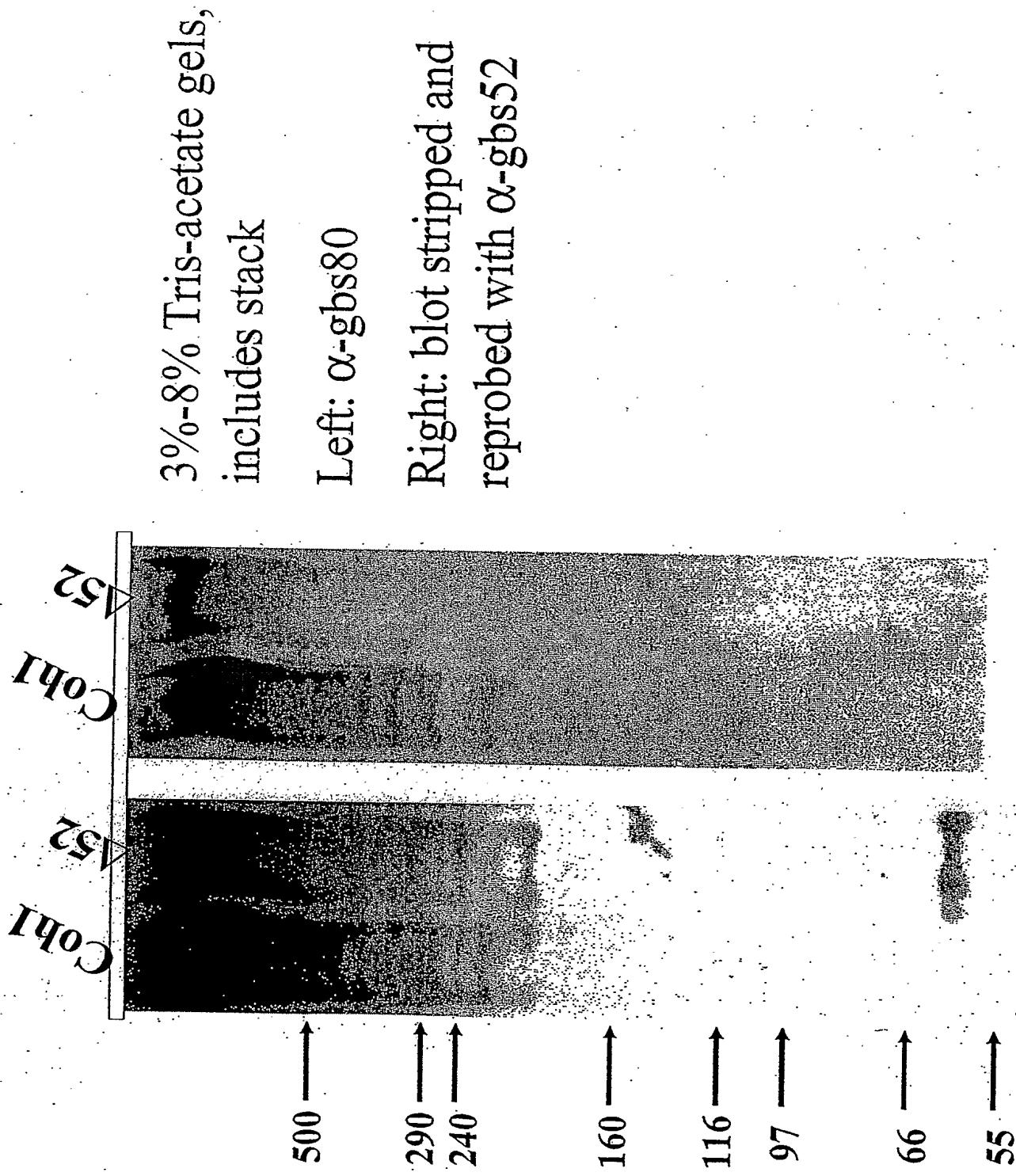


Figure 46: The pilus is found in the supernatant of the bacterial culture

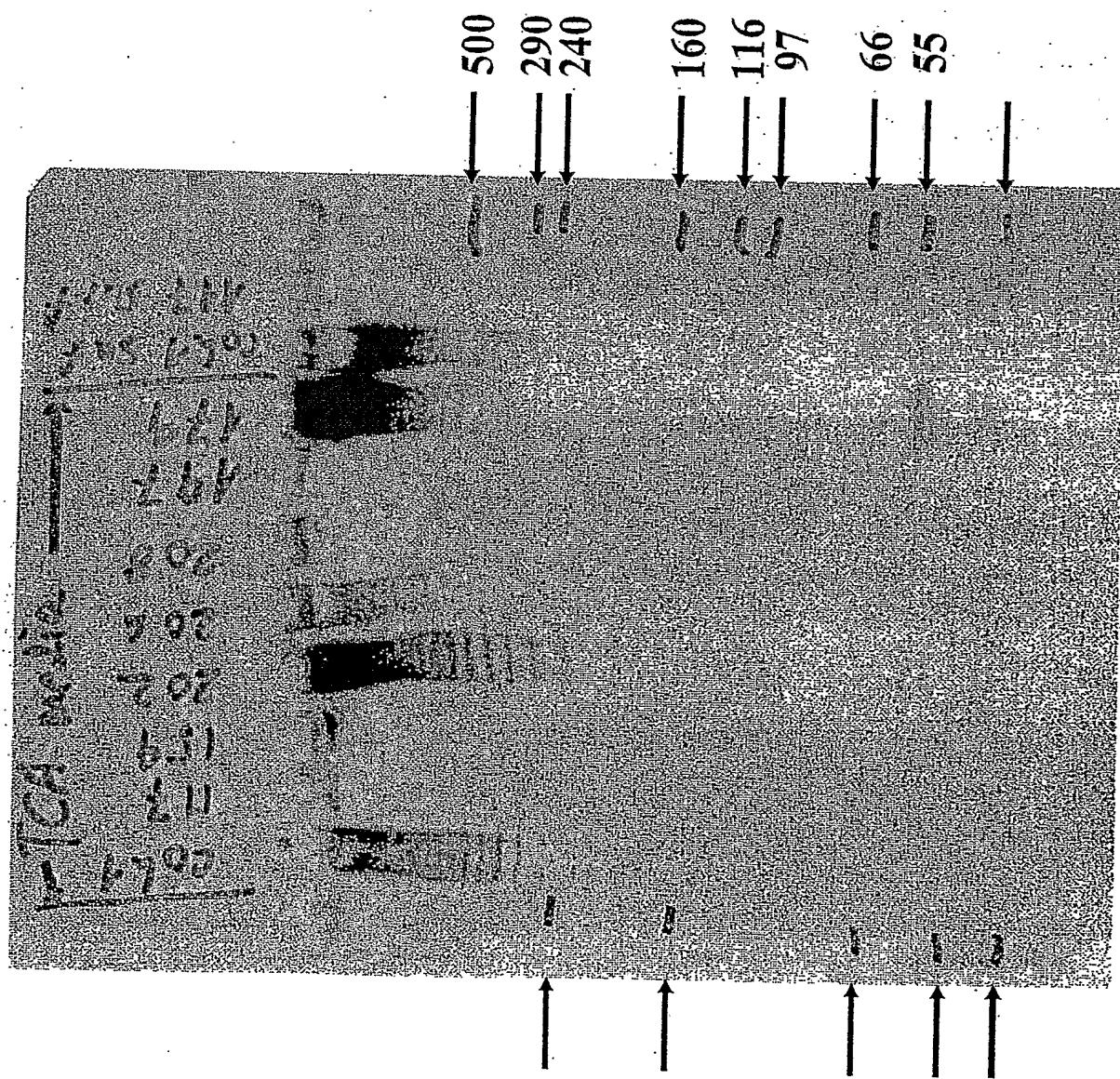


Figure 47: The pilus is found in the supernatant of cultures in all growth phases

TCA precipitation of 1 ml of THB culture supernatant run on 3-8% SDS-PAGE. OD600 nm are noted above samples, "F" indicates supernatant was filtered (0.2 μ M syringe filter).

Left five samples: Cohl.

Right five samples: 179
(Δ Gbs80/pGbs80).

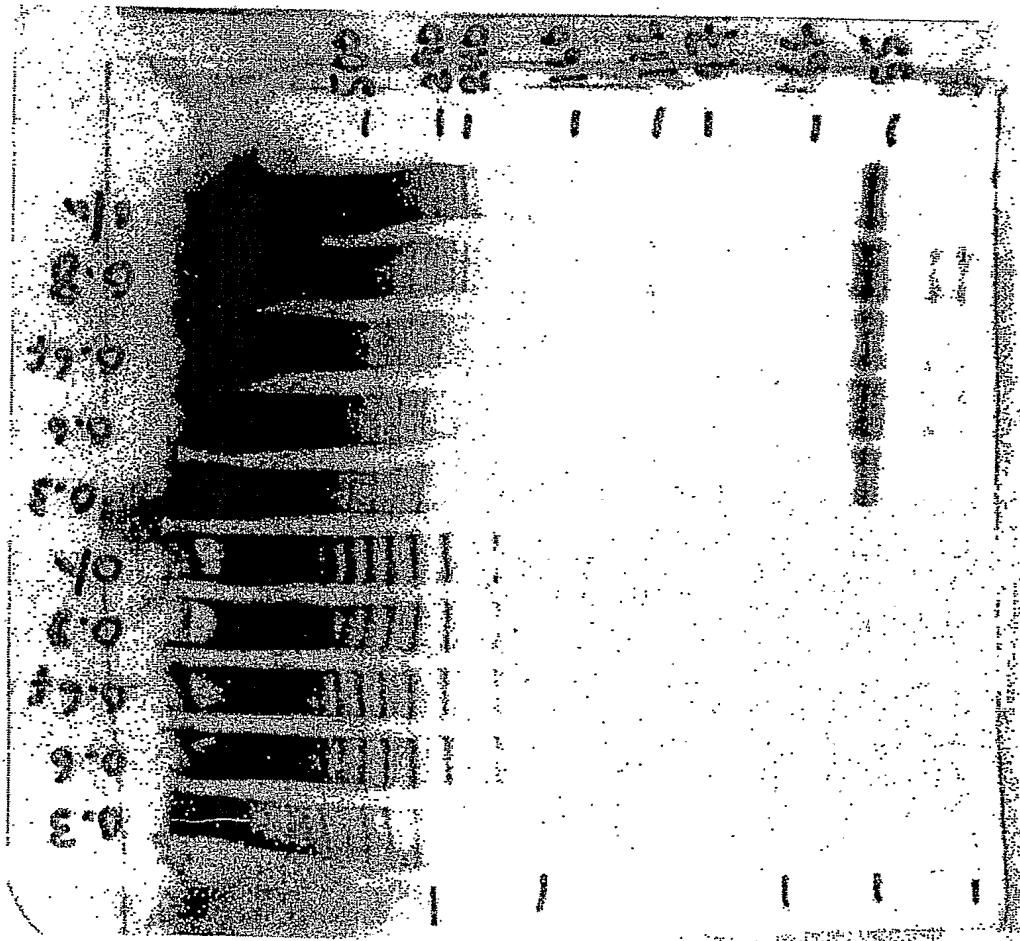


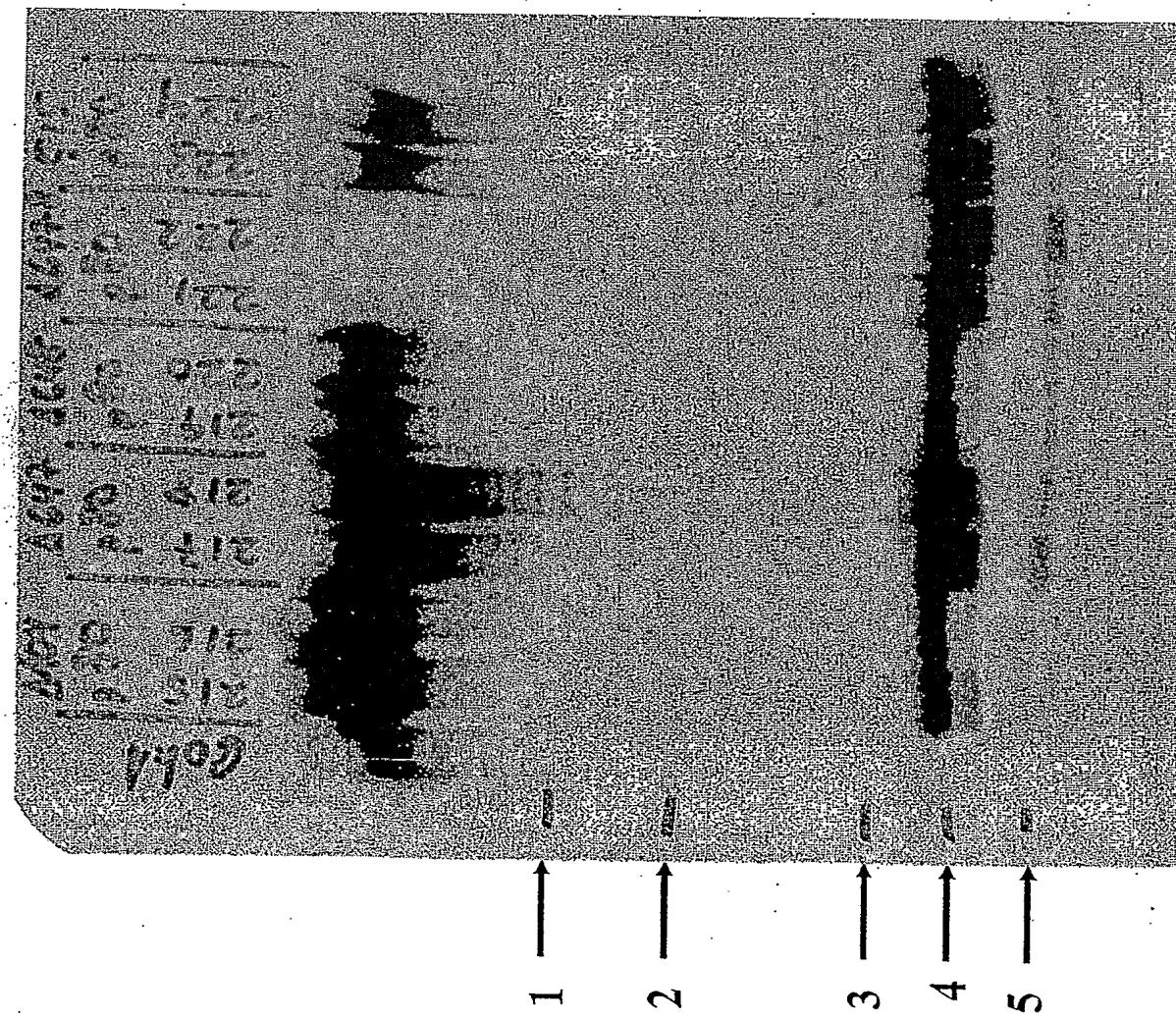
Figure 48: In *Coh1*, only the *gbs80* protein and one sortase (*sag0647* or *sag0648*) is required for polymerization

Over expression of *gbs80* in various strain backgrounds (two clones each).

Total protein extract preparations.

Only the double sortase mutant does not polymerize *gbs80*.

Gbs80 is polymerized in the *DK515* strain background (lacks adhesin island 1, adhesin island 2 is 2603-like). Presumably, *sag1405*&*sag1406* are responsible for polymerization.



GBS STRAIN JM9030013

IEM anti-GBS80

WO 2006/078318

PCT/US05/27239

164/487

PCT/US2005/027239

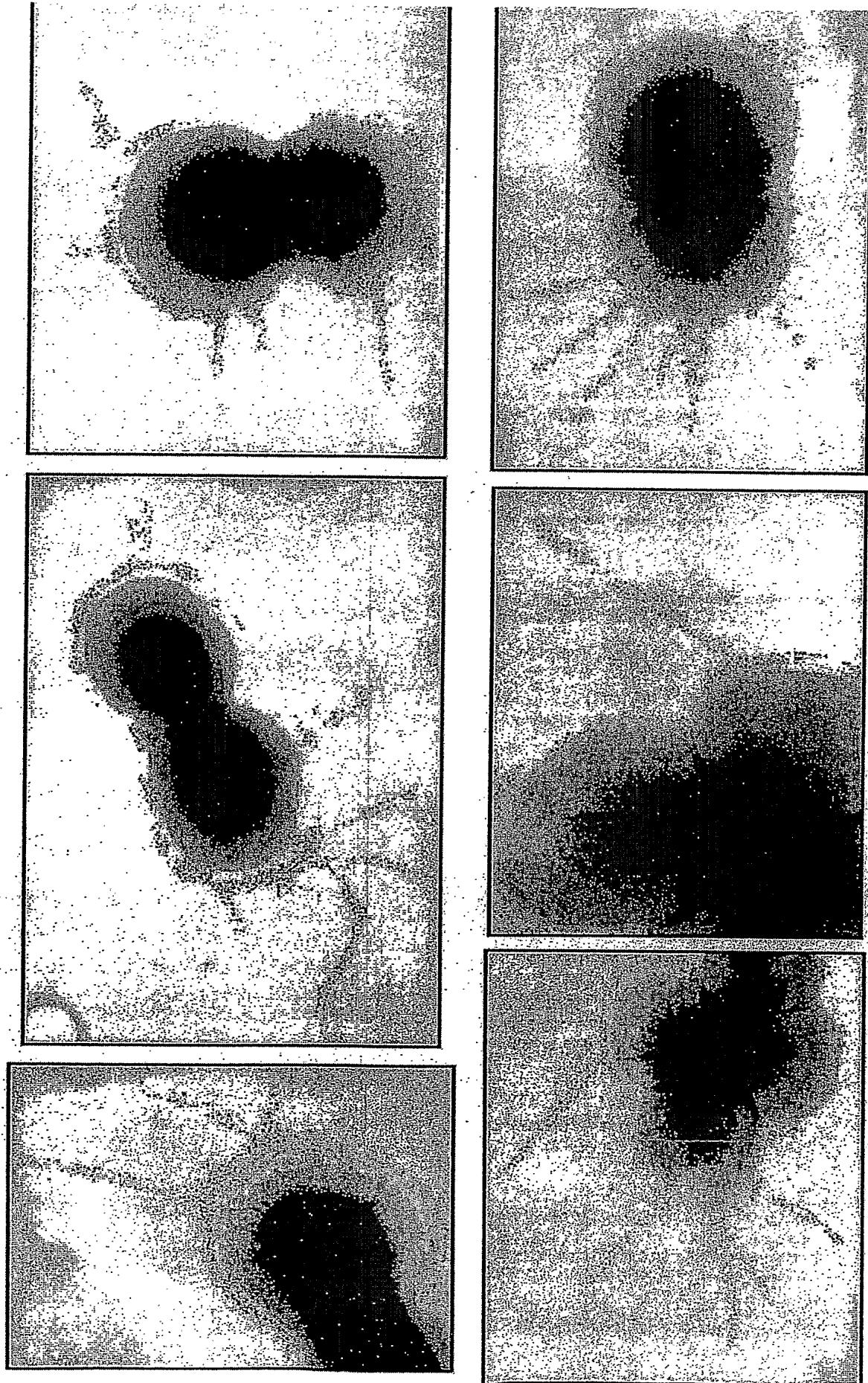
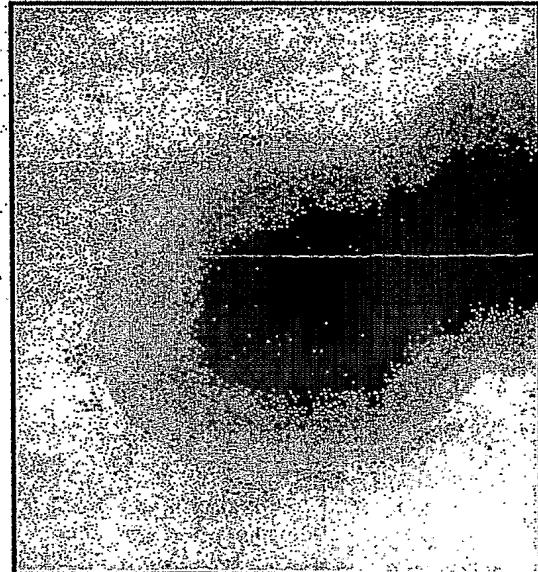
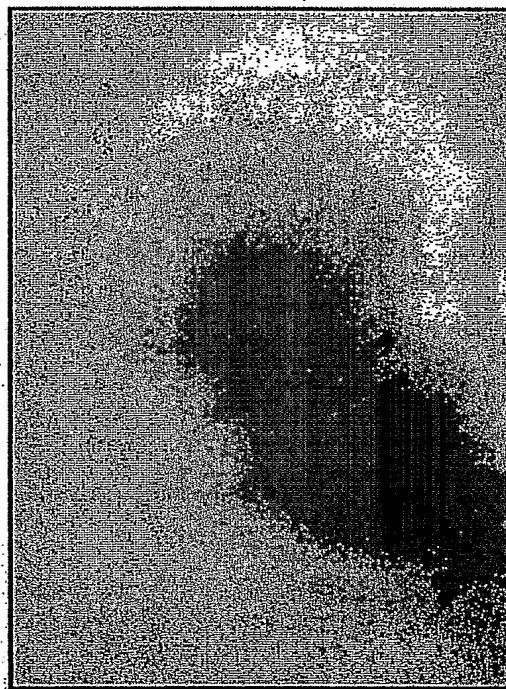
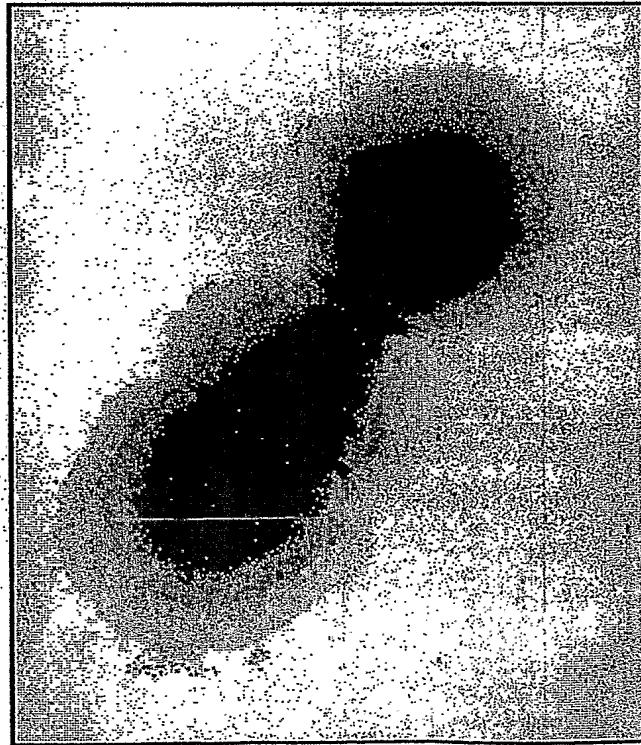


FIGURE 49

GBS STRAIN TM9030013

IEM anti-GBS104



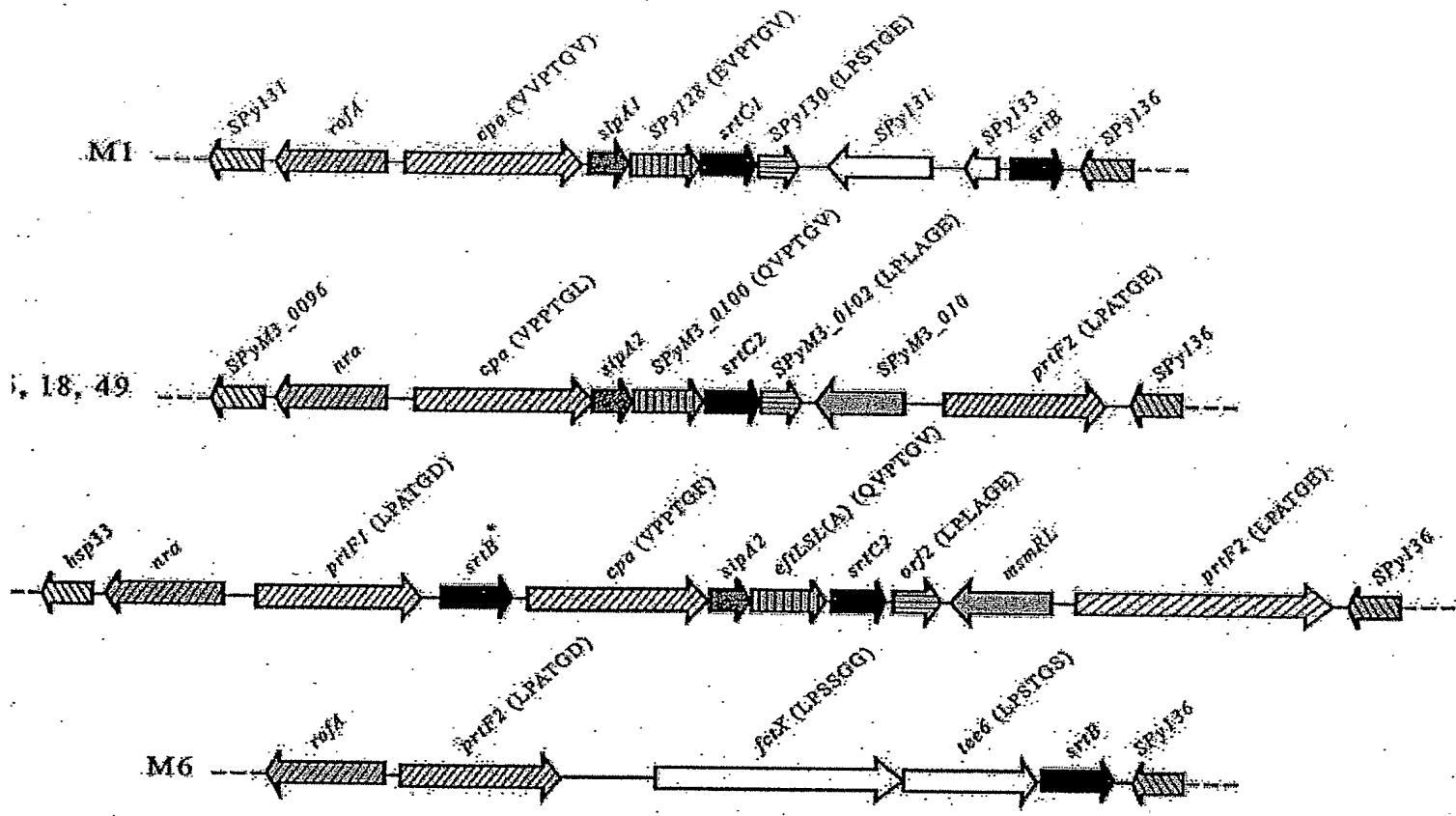
WO 2006/078318

PCT/US05/227239

165/487

PCT/US2005/027239

FIGURE 50.



KLEIN-FAG-IQCTDRAFTSLEENPIDIYDISSKNEEHNHKKOKKVCPRELLTNESEKVALSTCEGQHTCCTVTCGIE
KLTVTIFACLKTEAFDQLVPMNPAITPQDQGEPWVYI5KRSKCFPKVLTKEVTFVAFTSCENFSTDNRKVIVVGTION
KLTVTIFACLKTEAFDQLVPMNPAITPQDQGEPWVYI5KRSKCFPKVLTKEVTFVAFTSCENFSTDNRKVIVVGTION
KLTWTIFACLKTEAFDQLVPMNPAITPQDQGEPWVYI5KRSKCFPKVLTKEVTFVAFTSCENFSTDNRKVIVVGTION
KLTWTIFACLKTEAFDQLVPMNPAITPQDQGEPWVYI5KRSKCFPKVLTKEVTFVAFTSCENFSTDNRKVIVVGTION
KLTWTIFACLKTEAFDQLVPMNPAITPQDQGEPWVYI5KRSKCFPKVLTKEVTFVAFTSCENFSTDNRKVIVVGTION

FIGURE 51

PC T/US605 / 27239

GI-19224135	1	MNNKK EOKKQDAPR-VSNRHP	KOITTYLVGVFLMFTLSSMRGAQSIEGEEK
ORF78	1	----- EOKRDKTNYGSANNKR	ROTTIGELKVFLTFVALIDG-----IVG-----
GI-21909634	1	----- MOKRDKTNYGSANNKR	ROTTIGLLKVFLTFVALIDGIVGFSIRAFG
GI-28810257	1	----- MOKRDKTNYGSANNKR	ROTTIGLLKVFLTFVALIDGIVGFSIRAFG
GI-19745301	1	----- MOKRDKTNYGSANNKR	ROTTIGLLKVFLTFVALIDGIVGFSIRAFG
GAS15	1	1 LRGEKMKNTRFENKLNTINTORVLSKNSKRFWTVLGVFLMIFALMTSMVG EKTIVFG	
GI-19224135	53	R EEEVSVPKIKSPDD--AYPWYGDSYDSSHEYERFRVAHDLRVNLNGSKSYQVYCFNL	
ORF78	39	-----	
GI-21909634	46	- AEEQSVPNKQSSVQ--DYPWYGDSYSKGYEDYSPLKTYHNLKVNLDSKEYOAYCFNL	
GI-28810257	46	- AEEQSVPNKQSSVQ--DYPWYGDSYSKGYEDYSPLKTYHNLKVNLDSKEYOAYCFNL	
GI-19745301	46	- AEEOST-----	
GAS15	58	- LWESSTPNAINPDSSSEYRHYGYESYVRGHEYKQFRVAHDLRVNLLEGSRSYQVYCFNL	
GI-19224135	111	N SHYPNPKAVFSKQWENRVDGTGEVTNYEOTPKIRGESLANKELSIMYNAYPKNANGYM	
ORF78	39	-----	
GI-21909634	103	T KHFFSKSDSVRSQWYKILEGTNENPIKLADEKPRIEDGOLQONFLRLEYNGYPNDRNGIM	
GI-28810257	103	T KHFFSKSDSVRSQWYKILEGTNENPIKLADEKPRIEDGOLQONFLRLEYNGYPNDRNGIM	
GI-19745301	52	-----	
GAS15	117	KNAFFLCSDSSVKKIYKRPDGISTKEEDYPMSPRTGDELNOKTERAVMYNHPONANGIM	
GI-19224135	171	D KIEPLNAILVTOQAWWYSDSSYGN-IKTLWASELKEGKIDFEQVLMREAYSKLIISDD	
ORF78	39	----- FSTRAFG	
GI-21909634	163	K GIDPLNAILVTONATWYTDSSYISDTSKAFOQQETELKLDSSOOLMLRNWALKRLINPK	
GI-28810257	163	K GIDPLNAILVTONATWYTDSSYISDTSKAFOQQETELKLDSSOOLMLRNWALKRLINPK	
GI-19745301	52	-----	
GAS15	177	EGLLEPLNAIRVTOEAWWYSDNAPISNPDESFKRESESNLVTSQSLMRQALNQLIDEN	
GI-19224135	230	L EETSKNKLIPQGSKLNIFVPODNS-----VONLLSAEYVPESPPAPGQSEPEPPVQTNT	
ORF78	46	----- AEEKSTETNT	
GI-21909634	223	E VEISLPNOVEANYQLSIFQSSDNT-----FQNLLSAEYVPDTTPPKPG-----EEPPAKTENT	
GI-28810257	223	E VEISLPNOVEANYQLSIFQSSDNT-----FQNLLSAEYVPDTTPPKPG-----EEPPAKTENT	
GI-19745301	52	----- ETKT	
GAS15	237	LATKMPKOVEDDEQLSIFESEDINGDKYNKGYQNLLSGGIWTKPPTPGDPPMPPNQPOP	
GI-19224135	284	S VIIRKYAEVDYSKLLEGATLRLTCEDILDFQEKVFSNQTCGEKIELSNGTYTLTEISSP	
ORF78	57	S VIIRKYAEVDYSKLLEGATLRLTCEDILDFQEKVFSNQTCGEKIELSNGTYTLTEISSP	
GI-21909634	275	S VIIRKYAEVDYSKLLEGATLNLIAQIEGSGFQEKIFDSNISGEKVELENGTYLSELNEP	
GI-28810257	275	S VIIRKYAEVDYSKLLEGATLNLIAQIEGSGFQEKIFDSNISGEKVELENGTYLSELNEP	
GI-19745301	57	S VIIRKYAEVDYSKLLEGATLNLIAQIEGSGFQEKIFDSNISGEKVELENGTYLSELNEP	
GAS15	297	SVLIRKYAEVDYSKLLEGATLQLTGDNVNSFOARFSENDSERIELSDGTYTLTEISSP	
GI-19224135	344	E GYKIAEPIKFRWVNKKVFIWQKDGSOVEMENKEAEPYSVEAYSDMJDNSYINPETFTP	
ORF78	117	E GYKIAEPIKFRWVNKKVFIWQKDGSOVEMENKEAEPYSVEAYSDMJDNSYINPETFTP	
GI-21909634	335	Q GYGVAPITFKVAEAKVLIKKREGOFVENONKEIAEPYSVIAFNDFEETGYLSD-----DFNN	
GI-28810257	335	Q GYGVAPITFKVAEAKVLIKKREGOFVENONKEIAEPYSVIAFNDFEETGYLSD-----DFNN	
GI-19745301	117	Q GYGVIAEPIKFRWVNKKVFIWQKDGSOVEMENKEAEPYSVEAYSDMJDNSYINPETFTP	
GAS15	357	AGYSIAEPIKFRWVNKKVFIWQKDGSOVEMENKEAEPYSVEAYSDMJDNSYINPETFTP	
GI-19224135	404	I GKFYYAKNKKDKSSQVVYCFNAIDLHSPFESEDGGTIDPDLISTMKEYKYTHTAGSDLFKY	
ORF78	174	I YAKFYYAKNKKDKSSQVVYCFNAIDLHSPFESEDGGTIDPDLISTMKEYKYTHAGSDLFKY	
GI-21909634	393	I GKFYYAKNNTNGTINQVVYCFNAIDLHSPFESEDGGTIDPDLISTMKEYKYTHAGSDLFKY	
GI-28810257	393	I GKFYYAKNNTNGTINQVVYCFNAIDLHSPFESEDGGTIDPDLISTMKEYKYTHAGSDLFKY	
GI-19745301	177	I GKFYYAKNANGTSQVVYCFNAIDLHSPFESEDGGTIDPDLISTMKEYKYTHAGSDLFKY	
GAS15	413	YAKFYYAKNNGTSQVVYCFNAIDLHSPFESEDGGTIDPDLISTMKEYKYTHAGSDLFKY	

FIGURE 52

PCT/US605/27239

GI-19224135	464	SLRPRDTNPEDFLKHINKVIEKGYNNKGD--SYNGLTEQFRARTQLAIYYFTDSIDLKT
ORF78	233	SINTARDEDPEOLFLKHINKVIEKGYNNKGD--SYNGLTEQFRARTQLAIYYFTDSIDLKT
GI-21909634	453	AATPRDKDADEFLKHKIKKILDKGYKKKGD--TYKTLTEAOFRARTQLAIYYFTDSIDLKT
GI-28810257	453	AATPRDKDADEFLKHKIKKILDKGYKKKGD--TYKTLTEAOFRARTQLAIYYFTDSIDLKT
GI-19745301	237	SNINPRASTNDELLSÖVNIVLEKGYRDDST--TVANLTSVEFFRARTQLAIYYFTDSIDLKT
GAS15	472	TVKPRDTDPDIFLKHINKVIEKGYREKGQATEYSGLTETOLRAATQLAIYYFTDSIDLKT
GI-19224135	522	LKTYNNNGKVHGPFESMDEKTLAVTKELINYAOD-NSAPQLTNLDFFVPNNSKYOSLIGTO
ORF78	291	--TKDRILNDHGFGDMNDOTLGVAKKIVEYALS-DEDSKLTNLDFFVPNNSKYOSLIGTO
GI-21909634	511	LKTYNDNKGYHGFDKLDATLAVVHILITYAED-VTELEMTONLDFFVPNSSRYQSLIGTO
GI-28810257	511	LKTYNDNKGYHGFDKLDATLAVVHILITYAED-VTELEMTONLDFFVPNSSRYQSLIGTO
GI-19745301	295	LADY---HGFCALTEALNATKETVAYAEDRANLPNISNLDFYVPNSNKYOSLIGTO
GAS15	531	--KEKLKDYGFGDMNDSTLAVAKTLEYAOD-SNPPLTDLDFFIPNNNKYOSLIGTO
GI-19224135	581	VHPDDLVDVIRMEDNKQEVIPTVHSITVKKTWVGELGDNTRGFQFEIELBLNDKTGOFIVNT
ORF78	348	VHPDDLVDVIRMEDNKQEVIPTVHSITVKKTWVGELGDNTRGFQFEIELBLNDKTGOFIVNT
GI-21909634	570	VHNFELIDVISMEDNCPPIIPITHKLTISKTVTGTIADNKKEFNFIEIHLNSSDGQAISGT
GI-28810257	570	VHNFELIDVISMEDNCPPIIPITHKLTISKTVTGTIADNKKEFNFIEIHLNSSDGQAISGT
GI-19745301	349	VHPESLVDIIRMEDNCPPIIPITHKLTISKTVTGTIADNKKEFNFIEIHLNSSDGQAISGT
GAS15	587	WHFEDLVDEIRMEDNK-EVIPVTHMLTIRKTVTLAGCDRINKDFHFEIBLNNNKQELISQT
GI-19224135	641	LKTNNOGLVAKDGKYSENLNHGDTIRIEGLPTGYSYIJKETE-KDYIWTVDNKVSOEAQS
ORF78	408	LKTNNOGLVAKDGKYSENLNHGDTIRIEGLPTGYSYIJKETE-KDYIWTVDNKVSOEAQS
GI-21909634	630	YPTNSGELTVTDGKATEILNGESLIVEGLPSGYSYEIDETGASDYEVSVNGKNAFDGKA
GI-28810257	630	YPTNSGELTVTDGKATEILNGESLIVEGLPSGYSYEIDETGASDYEVSVNGKNAFDGKA
GI-19745301	409	YPTNSGELTVTDGKATEILNGESLIVEGLPSGYSYEIDETGASDYEVSVNGKNAFDGKA
GAS15	646	VKTDKTNLEFKDGKATINLNGESIYLOGLPEGYSYIJKETESEGYKVKAISOEVANATV
GI-19224135	701	ASENVITADKEVIFENRMDLVPPPTGEITDGGIYLULLLVLPGGLLWLFGRKGLND-
ORF78	468	ASENVITADKEVIFENRMDLVPPPTGLTDGAIYLULLLVLPGGLLWLFGRKGLND-
GI-21909634	690	TKASVKEDETVIFENRMDLVPPPTGLTDGAIYLULLLVLPGGLLWLFGRKGLND--
GI-28810257	690	TKASVKEDETVIFENRMDLVPPPTGLTDGAIYLULLLVLPGGLLWLFGRKGLND--
GI-19745301	469	TKASVKEDETVIFENRMDLVPPPTGLTDGAIYLULLLVLPGGLLWLFGRKGLND--
GAS15	706	SNTCITSDETLAFENNKEPVPTGVQDKINGYLALIVIACISLCHIGHTEFRKHD

FIGURE 52A

GI-1922WO 2006/078318 S_YM_FAR_EKMNNK_MI_LNKEAGFLVHTKRRKF_AVTLVGVF_FLLACAGAIGFGQVAY
 GI-50913503 S_YM_FAR_EKMNNK_MI_LNKEASFLAHTKRRKF_AVTLVGVF_FLLACAGAIGFGQVAY

GI-19224134 61 AADEKTVFNF_SPDPDYPMYGYDSY-----RGIFARYHNLKVNLKGSK_EYQAYCNETK
 GI-50913503 61 AADEKTVF_SHSSPNPEFPMYGYD_AGKEYPGYNIMTRYHDLRVNLNGSRSYQVYCNETK

GI-19224134 115 YFPRPTYSTTNNEKKIDGSGSAEKS_AANP_VLLENL_KLEKN_LLN_VYNGVKSNANGF
 GI-50913503 121 N_VPSQKN_SFIKKNWEKKIEGNCKSFVDYAH_TK_E-----ELEORELS_EYNEYPN_DANGY

GI-19224134 175 MNGI_EEDLNAILVT_NATI_YYSDS_APLNDVNKM_ERPVNGEISESQV_TLMREALNKLIDP
 GI-50913503 178 MKGLEE_BLN_AITVTQYAV_YYSDNS-QYQFETL_ES_AE_GK_ISRSQV_TLMREALNKLIDP

GI-19224134 235 NLEATAANKIPSGYRLNIFKSE_EYQ_NLLSAEYVPDDPPKPGDTSEHNPKTP_ELDGTPI
 GI-50913503 237 NLEATAVNKIPSGYRLNIF_ESENEAYQ_NLLSAEYVPDDPPKPG_ETS_EHNPKTP_ELDGTPI

GI-19224134 295 PEDPKRPDESS_EALPPLMP_ELDGEEVPEV_PSES_LE_PALPPLMP_ELDGEEVPEV_PSES_L
 GI-50913503 297 PEDPKHPDDN_EPTLFFVM-----

GI-19224134 355 PALPPLMP_ELDGEEVPEV_PSES_LE_PALPPLMP_ELDGEEVPEV_PSES_LE_PALPPLMP_ELDG
 GI-50913503 316 -----LDGEEVPEV_PSES_LE_PALPPLMP_ELDG

GI-19224134 415 BEVPEKPSV_DLPIEV_PRYEFNNKDQS_EPLAGESGETEYITEV_VGNQ_QNPV_DIDKKLPNETG
 GI-50913503 343 QEVPEKPS_DLPIEV_PRYEFNNKDQS_EPLAGESGETEYITEV_VGNQ_QNPV_DIDKKLPNETG

GI-19224134 475 FSGNNVETEDTKEPEV_LMGGQ_SES_EFTKDTOTGM_SGOTTPOVETEDTKEPEV_LMGGQ_SE
 GI-50913503 403 FSGNNVETEDTKEPEV_LMGGQ_SES_EFTKDTOTGM_SGOTTPOVETEDTKEPEV_LMGGQ_SE

GI-19224134 535 SVEFTKDTOTGM_SGOTTPOVETEDTKEPEV_LMGGQ_SES_EFTKDTOTGM_SGOTTPOVETE
 GI-50913503 463 SVEFTKDTOTGM_SGOTTPOVETEDTKEPEV_LMGGQ_SES_EFTKDTOTGM_SGOTTPOVETE

GI-19224134 595 DTKEE_VLNGGQ_SES_EFTKDTOTGM_SG_ESET_VT_VEDTRPKLVFHFDNNEPKVEENREK
 GI-50913503 523 DTKEE_VLNGGQ_SES_EFTKDTOTGM_SG_ESET_VT_VEDTRPKLVFHFDNNEPKVEENREK

GI-19224134 655 PTKNITPILPATGDIENVL_AFLGIL_LLSIFSLLNNK_RINKV-
 GI-50913503 583 PTKNITPILPATGDIENVL_AFLGIL_LLSIFSLLNNK_ROSINKV-

170/487

GI-19745307 1 MTQKNSY FLLSLTGFI1CGLLVTIGLSGVVGHAETRNGANKOGA PCT/US2005/027239
ORF84 WO 2006/078318 INSY-SFLSLTGFI1CGLLVTIGLSGVVGHAETRNGANKOGA
GI-28810263 1 MTQKNSYKLSFLLSLTGFI1CGLLVTIGLSGVVGHAETRNGANKOGA FEINKN
GI-21909640 1 MTQKNSYKLSFLLSLTGFI1CGLLVTIGLSGVVGHAETRNGANKOGA FEINKN
GI-19224141 1 MTQKNSYKLSFLLSLTGFI1CGLLVTIGLSGVVGHAETRNGANKOGSFEINKN DQNNKP

GI-19745307 55 ----- KSQEENYE
ORF84 55 ----- KSQEENYE
GI-28810263 55 ----- KSQEENYE
GI-21909640 1 -----
GI-19224141 61 LPGATFSLTSKDKGTSVQFTSNDKGIVDAQNLQPGTYTLKEETAPDGYD RISRTWTVT

GI-19745307 64 VYDN ----- RNI
ORF84 64 VYDN ----- RNI
GI-28810263 64 VYDN ----- RNI
GI-21909640 1 -----
GI-19224141 121 VYENGYTKLVENPYNGEIISKAGSKDVSSQLQLENPKMSVVS SKY GKT EVSSGAADFY RNH

GI-19745307 71 LQDGHEHKLEIKRVDGTGKTYQG FCFOLTKNFP TAQGVSKLYKKLSS
ORF84 71 LQDGHEHKLEIKRVDGTGKTYQG FCFOLTMIFP TAQGVSKLYKKLSS
GI-28810263 71 LQDGHEHKLEIKRVDGTGKTYQG FCFOLTKNFP TAQGVSKLYKKLSS
GI-21909640 1 ----- MSS
GI-19224141 181 AAYFKMSF EIKQDKSETINP CDTFVLQIDRRLNPKGIS QDHPKIEYDSANSPLAIGKYH

GI-19745307 118 ----- SDEETLK
ORF84 118 ----- SDEETLK
GI-28810263 118 ----- SDEETLK
GI-21909640 4 ----- SDEETLK
GI-19224141 241 AENHQLIYTFTDYIAGLDKVQLSAELSLFLEMKEVLENTSISNFKSTIGGQEITYKGTVN

GI-19745307 125 ----- QYASKYTSNRRGDTSG
ORF84 125 ----- QYASKYTSNRRGDTSG
GI-28810263 125 ----- QYASKYTSNRRGDTSG
GI-21909640 11 ----- QYASKYTSNRRGDTSG
GI-19224141 301 VLYGNESTKESNYITNGLSVCGSIESYNTETGEFVWYVYVNPNRNTNIPYATMNLWGFGF

GI-19745307 141 ----- NLRKQIANVLTEGYPT
ORF84 141 ----- NLRKQIANVLTEGYPT
GI-28810263 141 ----- NLRKQIANVLTEGYPT
GI-21909640 27 ----- NLRKQIANVLTEGYPT
GI-19224141 361 ARSNTSDILENDANTSSAELGEIQVYEVPEGEKLPSYGVDTVKTLMLEQDITAGLCNGCFOM

GI-19745307 157 NKS DMLN ----- GLTENEKIEVTQDAIWYF
ORF84 157 NKS DMLN ----- GLTENEKIEVTQDAIWYF
GI-28810263 157 NKS DMLN ----- GLTENEKIEVTQDAIWYF
GI-21909640 43 NKS DMLN ----- GLTENEKIEVTQDAIWYF
GI-19224141 421 TFRQRIDEGNNTQNKAFI1KVTGKTDQSGKPLVQSNLASFRCASEYAAFTPVGGMWFO

GI-19745307 182 TETTVPADR ----- SYTNRNVNSQNMKEVYOKLIDTTDID KYEDVQFDLFVPQDTN
ORF84 182 TETTVPADR ----- SYTNRNVNSQNMKEVYOKLIDTTDID KYEDVQFDLFVPQDTN
GI-28810263 182 TETTVPADR ----- SYTNRNVNSQNMKEVYOKLIDTTDID KYEDVQFDLFVPQDTN
GI-21909640 68 TETTVPADR ----- SYTNRNVNSQNMKEVYOKLIDTTDID KYEDVQFDLFVPQDTN
GI-19224141 481 NEIAALSPSKGSGSGKSEFTKPSITVANLVRQLRFKKMSTDNVPLPEAAFELRSSNGNS

GI-19745307 233 -- LQAVISVEPVIESLPVTS ----- LRPPIAQKDITAKN
ORF84 233 -- LQAVISVEPVIESLPVTS ----- LRPPIAQKDITAKN
GI-28810263 233 -- LQAVISVEPVIESLPVTS ----- LRPPIAQKDITAKN
GI-21909640 119 -- LQAVISVEPVIESLPVTS ----- LRPPIAQKDITAKN
GI-19224141 541 QKLEASSNTQGEWHFKDLTS GTYDLYETKAPKGYQQVTEKEETVTVDTTPAEEMVTWGS

FIGURE 54

GI-19745307 264 ----- TWW KERKITYFKLYRQLGEKEVAVDDAELK ----- QINS
 ORF84 WO 2006/078318 TWW KERKITYFKLYRQLGEKEVAVDDAELK ----- QINSEGOPCT/US2005/027239
 GI-28810263 264 ----- TWDAPKERKITYFKLYRQLGEKEVAVDDAELK ----- QINSEGQ
 GI-21909640 150 ----- TWDAPKERKITYFKLYRQLGEKEVAVDDAELK ----- QINSEGQ
 GI-19224141 601 PHSSVKVEANKEVTIMNHNETLTFSGKNIWENDREDORPAKIQVQLQNGOKMPNQIQEV

GI-19745307 305 ----- QEISVTVTNQLVTDEKGMRITYISVKEVDK ----- NGELLEPKD ----- YIKKED
 ORF84 305 ----- QEISVTVTNQLVTDEKGMRITYISVKEVDK ----- NGELLEPKD ----- YIKKED
 GI-28810263 305 ----- QEISVTVTNQLVTDEKGMRITYISVKEVDK ----- NGELLEPKD ----- YIKKED
 GI-21909640 191 ----- QEISVTVTNQLVTDEKGMRITYISVKEVDK ----- NGELLEPKD ----- YIKKED
 GI-19224141 661 TKDNEWSYHEKDLPKYDANNQEYKYSVEVNVPDGYKVSYLGNEFTNTRETEFVEHQNNF

GI-19745307 349 GLTVTNTYVKPTSG HYDIEVTFC ----- NGHIDITEDTPDIVSGENQMK
 ORF84 349 GLTVTNTYVKPTSG HYDIEVTFC ----- NGHIDITEDTPDIVSGENQMK
 GI-28810263 349 GLTVTNTYVKPTSG HYDIEVTFC ----- NGHIDITEDTPDIVSGENQMK
 GI-21909640 235 GLTVTNTYVKPTSG HYDIEVTFC ----- NGHIDITEDTPDIVSGENQMK
 GI-19224141 721 NLEFGNAEINGQSGSKIIDEDTLTSFKGKKIWKNDAENRPQAIVQOLYADGVAVEGOTK

GI-19745307 394 QIEGEDS ----- KPIDEVT ----- ENNLIEFGKNTMPGEE
 ORF84 394 QIEGEDS ----- KPIDEVT ----- ENNLIEFGKNTMPGEE
 GI-28810263 394 QIEGEDS ----- KPIDEVT ----- ENNLIEFGKNTMPGEE
 GI-21909640 280 QIEGEDS ----- KPIDEVT ----- ENNLIEFGKNTMPGEE
 GI-19224141 781 EISGSGNEWSFEFKNLKKYNGTGNDIIYSVKEVTVPGYDVTTYSANDITNTKREVITQQG

GI-19745307 424 ----- DGTNSNKYEEVEDSRPVDTLSGLSSEQGQSGDMTIEEDSATIHKFSKRD
 ORF84 424 ----- DGTNSNKYEEVEDSRPVDTLSGLSSEQGQSGDMTIEEDSATIHKFSKRD
 GI-28810263 424 ----- DGTNSNKYEEVEDSRPVDTLSGLSSEQGQSGDMTIEEDSATIHKFSKRD
 GI-21909640 310 ----- DGTNSNKYEEVEDSRPVDTLSGLSSEQGQSGDMTIEEDSATIHKFSKRD
 GI-19224141 841 PKLEIEETLPLESGASGGTTIEDSRPVDTLSGLSSEQGQSGDMTIEEDSATIHKFSKRD

GI-19745307 473 IDGKELAGATMELRDSGGNTISTWISDGQVDFYLMPGKYTFVETAAPDGYEIATAITFT
 ORF84 473 IDGKELAGATMELRDSGGNTISTWISDGQVDFYLMPGKYTFVETAAPDGYEIATAITFT
 GI-28810263 473 IDGKELAGATMELRDSGGNTISTWISDGQVDFYLMPGKYTFVETAAPDGYEIATAITFT
 GI-21909640 359 IDGKELAGATMELRDSGGNTISTWISDGQVDFYLMPGKYTFVETAAPDGYEIATAITFT
 GI-19224141 901 IDGKELAGATMELRDSGGNTISTWISDGQVDFYLMPGKYTFVETAAPDGYEIATAITFT

GI-19745307 533 VNEQGQVTVNGKATKGDASHIVIVDAYKPTKGSGQVIDIEEKLPEOGHSGSTTEIEDSKS
 ORF84 533 VNEQGQVTVNGKATKGDASHIVIVDAYKPTKGSGQVIDIEEKLPEOGHSGSTTEIEDSKS
 GI-28810263 533 VNEQGQVTVNGKATKGDASHIVIVDAYKPTKGSGQVIDIEEKLPEOGHSGSTTEIEDSKS
 GI-21909640 419 VNEQGQVTVNGKATKGDASHIVIVDAYKPTKGSGQVIDIEEKLPEOGHSGSTTEIEDSKS
 GI-19224141 961 VNEQGQVTVNGKATKGDASHIVIVDAYKPTKGSGQVIDIEEKLPEOGHSGSTTEIEDSKS

GI-19745307 593 SDVIIGGQG -----
 ORF84 593 SDVIIGGQG -----
 GI-28810263 593 SDVIIGGQGEVVDTTEDTOSGMTGHS -----
 GI-21909640 479 SDVIIGGQGEVVDTTEDTOSGMTGHSSTTEIEDSKSSDVIIIGQGEVVDTTEDTOSGMT
 GI-19224141 1021 SDLIIGGQGEVVDTTEDTOSGMTGHS -----

GI-19745307 602 ----- DIVETTEDTGTGMHDSGCITEVEDTKLVQSFHF DNK
 ORF84 602 ----- DIVETTEDTGTGMHDSGCITEVEDTKLVQSFHF DNK
 GI-28810263 619 ----- GSTTKEEDSKSSDVIVCGQGQIVETTEDTGTGMHDSGCITEVEDTKLVQSFHF DNK
 GI-21909640 539 GHSGSTTKEEDSKSSDVIVCGQGQIVETTEDTGTGMHDSGCITEVEDTKLVQSFHF DNK
 GI-19224141 1047 ----- GSTTEIEDSKSSDVIIIGQGQIVETTEDTGTGMHDSGCITEVEDTKLVQSFHF DNK

GI-19745307 639 ESESNSEIPKKDKPKSNTSLPATGEKQHNMFIVMTCSLTSSVFVISLTKRLSSC
 ORF84 639 ESESNSEIPKKDKPKSNTSLPATGEKQHNMFIVMTCSLTSSVFVISLTKRLSSC
 GI-28810263 676 EPESENSEIPKKDKPKSNTSLPATGEKQHNMFIVMTCSLTSSVFVISLTKRLSSC
 GI-21909640 599 EPESENSEIPKKDKPKSNTSLPATGEKQHNMFIVMTCSLTSSVFVISLKSRLSSC
 GI-19224141 1104 EPESENSEIPKKDKPKSNTSLPATGEKQHNMFIVMTCSLTSSVFVISLKSRLSSC

FIGURE 54A

GI-19224WO 2006/078318-MKRNKLLATAILATALGTASLNQNVKAETAGVUSSGOLTIKNSITPCT/US2005/027239

ORF80 PCT¹ 1 MERERMKRNKLLATAILATALGTASLNQNVKAETAGVUSSGOLTIKNSITPCT/US2005/027239
 GI-21909636 1 MERERMKRNKLLATAILATALGTASLNQNVKAETAGVUSSGOLTIKNSITPCT/US2005/027239
 GI-28810259 1 MERERMKRNKLLATAILATALGTASLNQNVKAETAGVUSSGOLTIKNSITPCT/US2005/027239
 GI-19745303 1 MERERMKRNKLLATAILATALGTASLNQNVKAETAGVUSSGOLTIKNSITPCT/US2005/027239
 GI-13621428 1 MERERMKRNKLLATAILATALGTASLNQNVKAETAGVUSSGOLTIKNSITPCT/US2005/027239

GI-19224137 56 PNTDYTFSVNPDSAPTCESN-LETRPGTAVN-NODIKVSYNTDKTSQNEKQVVADFMK
 ORF80 60 PTAETFTIIPDMTASGREGS-LEIINNGIVEGLDKQVIVKYKNTDKTSQNTKIAQDFDFSK
 GI-21909636 56 PRADYTFKVEADSTASGKTKDGLEIKPGIVNGLT-EQIISYTNTDNDPSIVKSTEFDPSK
 GI-28810259 61 PRADYTFKVEADSTASGKTKDGLEIKPGIVNGLT-EQIISYTNTDNDPSIVKSTEFDPSK
 GI-19745303 56 PRADYTFKVEADDNAKGKTKDGLEIKPGIVDGLENKTIIHYCNSDKTTANEKSVNFDFAN
 GI-13621428 48 PNTDFTFNIIEPDITVNEDGNKFKG-----VALNTPMTKVITYTNSDNGGSNTKTAEFDSE

GI-19224137 114 VTFPSVGTIYRYVVTENKGTAE-GVTYDDTKILVDVYVGN--NEKGCGLEPKYIVSNKEDSA
 ORF80 119 VKFFEAIGVYRYMVEKNDKD-GITYCDKKITVDVYVGNKANNEBGFELVYIVSKECTSS
 GI-21909636 115 VVFPGIGVYRYTVSEKQGDE-GITYDTKKIITVDVYVGN--KEGGGFEPKIEIVSKEDQTD
 GI-28810259 120 VVFPGIGVYRYTVSEKQGDE-GITYDTKKWTVDVYVGN--KEGGGFEPKIEIVSKEDQTD
 GI-19745303 116 VKFPQGVYVRYTVSEVNGNNA-GIAYDQQQNTVDVYVGN--REDGGFEAKYIVSTECCQS
 GI-13621428 103 VTEEKPGVYVYKVTEEKIDKVPGVSYDTTSVTVQHILWLN-EFQQKPVATIYIVGYKEGS

GI-19224137 171 TMEPIOFNNSFETTSLSKIEEVGTGNTGDKHAFTFTLTLQNEYEAASSVVKIEENCO--
 ORF80 178 TRKPKIPTFNSLKKTSLKIEKOITGNAQDRKASFNFTLTLQSEYNTGSVVKIEQDGS--
 GI-21909636 172 VNKPKVNFNNNSFATTSLKVKKNVSGNTGELQKEFDFTLTLQNESTNEKNDQIVSLOKGN--
 GI-28810259 177 VNKPKVNFNNNSFATTSLKVKKNVSGNTGELQKEFDFTLTLQNESTNEKNDQIVSLOKGN--
 GI-19745303 173 DKPKPTEKNEFDTTSLSKVKRVTGNTGEHORSPSFTLIIPNECHENGQVNILOGGB--
 GI-13621428 161 KVPIOKNSLSDSTTITVKKVSGTGCDRSKDFNFCITLKANQYXRASEKWMIENTTKGG

GI-19224137 229 ----TKDVKIGEAVKETLNDOSVILSKLFVGINYKVBEAEANOGGYTTATLNDG--EK
 ORF80 236 ----KKDWIGTPIKFTLGKGSVMSKLPIGINYVMSSEDEANKDGYTTATLNEOGKEK
 GI-21909636 230 ----KEVKIGTPYKEELKNGESTGDLKLPVGIVYKVBEAEANNDGYTTASLNEG-DGO
 GI-28810259 235 ----KEVKIGTPYKEELKNGESTGDLKLPVGIVYKVBEAEANNDGYTTASLNEG-DGO
 GI-19745303 231 ----TAKVIGEPPFTLKDQESVILSKLPVGIVYRVTEEDVTHDGWTSETLNDG-E-V
 GI-13621428 220 QAPVQTEASTDOEVHFTLKDGESIKVTLNLPGVDTWVTEDDYKSEKTTINVEVSPQDGAV

GI-19224137 283 LSTVNLG-OEHKTDKTADEIVVTNNRDTQVPTGVVGT LAPFAVL SIVAIGGVIVITKRN
 ORF80 292 SSEFPLSTONOKTDESAD EIVVTNKRD TOVPTGVVGT LAPFAVL SIVAIGGVIVITKRN
 GI-21909636 285 SKMYOOLD-MEOKTDESAD EIVVTNKRD TOVPTGVVGT LAPFAVL SIVAIGGVIVITKRN
 GI-28810259 290 SKMYOOLD-MEOKTDESAD EIVVTNKRD TOVPTGVVGT LAPFAVL SIVAIGGVIVITKRN
 GI-19745303 285 TDCYMLG-DSKQITDKSDEIVVTNKRD TOVPTGVVGT LAPFAVL SIVAIGGVIVITKRN
 GI-13621428 280 KNTAGNSTEQETSTDKDMTIPETNKKDIEVPTGVAMTVAPYALGIVAVGCALYEVNNKH

GI-19224137 342 A
 ORF80 352 A
 GI-21909636 344 A
 GI-28810259 349 A
 GI-19745303 344 A
 GI-13621428 340 A

173/487

PCT/US2005/027239

GI-2190WO 2006/078318

GI-28810261	1	MLFSVVMILTMALAFNQTVLAKDSTV
GI-19224139	1	MLFSVVMILTMALAFNQTVLAKDSTV
ORF82	1	LLFQRVKIFLLTIVLSLSVLFKNNERRRLRKYWKMLFSVVMILTMALAFNQTVLAKDSTV
GI-19745305	1	MRKYWKMLFSVVMILTMALAFNQTVLAKDSTV

GI-21909638	20	QTSISVENVLERAGDSTPFSTIALESIDAMKTIIEITIAGSGKASFSPLTFTTVGQTYRV
GI-28810261	26	QTSISVENVLERAGDSTPFSTIALESIDAMKTIIEITIAGSGKASFSPLTFTTVGQTYRV
GI-19224139	26	QTSISVENVLERAGDSTPFSTIALESIDAMKTIIEITIAGSGKASFSPLTFTTVGQTYRV
ORF82	61	QTSISVENVLERAGDSTPFSTIALESIDAMKTIIEITIAGSGKASFSPLTFTTVGQTYRV
GI-19745305	32	QTSISVENVLERAGDSTPFSTIALESIDAMKTIIEITIAGSGKASFSPLTFTTVGQTYRV

GI-21909638	80	YQKPSQNKDQADTTVFDVLVYVTYDEDGTLVAKVISRRAGDEEKSAITFKPKRNLVKPIP
GI-28810261	86	YQKPSQNKDQADTTVFDVLVYVTYDEDGTLVAKVISRRAGDEEKSAITFKPKRNLVKPIP
GI-19224139	86	YQKPSQNKDQADTTVFDVLVYVTYDEDGTLVAKVISRRAGDEEKSAITFKPKRNLVKPIP
ORF82	121	YQKPSQNKDQADTTVFDVLVYVTYDEDGTLVAKVISRRAGDEEKSAITFKPKRNLVKPIP
GI-19745305	92	YQKPSQNKDQADTTVFDVLVYVTYDEDGTLVAKVISRRAGDEEKSAITFKPKRNLVKPIP

GI-21909638	140	PRQPNIPKTPLEPLAGEVKSLLGILSIVLLGLLVLLYVKKLKSRL
GI-28810261	146	PRQPNIPKTPLEPLAGEVKSLLGILSIVLLGLLVLLYVKKLKSRL
GI-19224139	146	PRQPNIPKTPLEPLAGEVKSLLGILSIVLLGLLVLLYVKKLKSRL
ORF82	181	PRQPNIPKTPLEPLAGEVKSLLGILSIVLLGLLVLLYVKKLKSRL
GI-19745305	152	PRQPDIPKTPLEPLAGEVKSLLGILSIVLLGLLVLLYVKKLKSRL

FIGURE 56

PCT/US2005/027239 174/487

!!SEQUENCE_LIST 1.0

(Peptide) FASTA of: gi-50913505.pep from: 1 to: 1036 September 15, 2004 18:46

gi|50913505|ref|YP_059477.1| Collagen adhesion protein [Streptococcus pyogenes M GAS10394]

TO: *.pep Sequences: 56 Symbols: 22,803 Word Size: 2

Scoring matrix: GenRunData:blosum50.cmp

Variable pamfactor used

Gap creation penalty: 12 Gap extension penalty: 2

Histogram Key:

Each histogram symbol represents 1 search set sequences

z-scores computed from opt scores

z-score	obs (=)	exp (*)
< 20	0	0:
22	0	0:
24	0	0:
26	0	0:
28	0	0:
30	0	0:
32	0	0:
34	0	1:*
36	0	2: *
38	0	3: *
40	0	4: *
42	0	5: *
44	0	5: *
46	5	5:====*
48	10	5:====*=====
50	12	4:====*=====
52	4	4:====*
54	6	3:====*====
56	4	3:====*=
58	5	2:==*====
60	3	2:==*==
62	0	1:*
64	1	1:*
66	1	1:*
68	0	1:*
70	1	1:*
72	0	0:
74	0	0:
76	1	0:=
78	0	0:
80	0	0:
82	0	0:
84	0	0:
86	0	0:
88	0	0:
90	0	0:
92	0	0:
94	0	0:
96	1	0:=
98	0	0:

FIGURE 57

PCT/US2005/027239
 100 0 0:
 102 0 0:
 104 0 0:
 106 1 0:=
 108 0 0:
 110 0 0:
 112 0 0:
 114 0 0:
 116 0 0:
 118 0 0:
 >120 1 0:=

Joining threshold: 39, opt. threshold: 27, opt. width: 16, reg.-scaled

The best scores are:

initl initn opt z-sc E(54)..

/home/morama/gas/pili/align/gi-50913505.pep		Begin: 1 End: 1036	
! gi 50913505 ref YP_059477.1 Collag...	6697	6697 3452.1	9.6e-189
/home/morama/gas/pili/align/gi-19224141.pep		Begin: 48 End: 144	
! gi 19224141 gb AAL86412.1 AF447492...	63	100 159 105.9	0.023
/home/morama/gas/pili/align/gi-21909640.pep		Begin: 147 End: 449	
! gi 21909640 ref NP_663908.1 protei...	35	35 136 96.3	0.08
/home/morama/gas/pili/align/gi-13621428.pep		Begin: 57 End: 318	
! gi 13621428 gb AAK33238.1 hypothet...	33	33 91 75.6	1.1
/home/morama/gas/pili/align/gi-50913506.pep		Begin: 33 End: 428	
! gi 50913506 ref YP_059478.1 Fimbri...	70	149 86 71.3	1.9
/home/morama/gas/pili/align/gi-13621432.pep		Begin: 14 End: 56	
! gi 13621432 gb AAK33241.1 conserve...	40	65 78 68.0	2.9
/home/morama/gas/pili/align/gi-19745301.pep		Begin: 241 End: 466	
! gi 19745301 ref NP_606437.1 putati...	52	52 73 64.8	4.3
/home/morama/gas/pili/align/gas15.pep		Begin: 492 End: 739	
! GAS15 GAS15	43	68 69 61.4	6.6
/home/morama/gas/pili/align/gi-21909636.pep		Begin: 176 End: 298	
! gi 21909636 ref NP_663904.1 conser...	31	31 62 60.8	7.1
/home/morama/gas/pili/align/gi-28810259.pep		Begin: 181 End: 303	
! gi 28810259 gb BAC63197.1 hypoth...	31	31 62 60.7	7.2
/home/morama/gas/pili/align/gi-19224139.pep		Begin: 90 End: 143	
! gi 19224139 gb AAL86410.1 AF447492...	43	43 54 58.9	8.9
/home/morama/gas/pili/align/gi-19745305.pep		Begin: 96 End: 149	
! gi 19745305 ref NP_606441.1 hypoth...	43	43 54 58.8	9
/home/morama/gas/pili/align/orf82.pep		Begin: 125 End: 178	
! TRANSLATE of: orf82.seq check: 4296...	43	43 54 58.2	9.6
/home/morama/gas/pili/align/gi-21909638.pep		Begin: 84 End: 137	
! gi 21909638 ref NP_663906.1 hypoth...	43	43 52 58.0	9.9
\End of List			

gi-50913505.pep

/home/morama/gas/pili/align/gi-50913505.pep

gi|50913505|ref|YP_059477.1| Collagen adhesion protein [Streptococcus pyogenes M GAS10394]

SCORES Init1: 6697 Initn: 6697 Opt: 6697 z-score: 3452.1 E(): 9.6e-189
 >>/home/morama/gas/pili/align/gi-50913505.pep (1036 aa)
 initn: 6697 init1: 6697 opt: 6697 Z-score: 3452.1 expect(): 9.6e-189
 Smith-Waterman score: 6697; 100.0% identity in 1036 aa overlap
 (1-1036:1-1036)

10 20 30 40 50 60
 gi-50913505. MYSRLKRELVIVINRKKYKLIRLMVTVGQLIFSQLVLPPIRRLGLQMISTQTKVIPQEIVT

FIGURE 57A

PC T11H99E51P7P3P

gi-50913505. MYSRLKRELVIVINRKKYKLIRLMVTVGLIFSQLVLPIRRGLQMI
10 20 30 40 50 60

gi-50913505. QTETQGTQVVATKQKLESENSSLKVALKRESGFEHNATIDASLDTE
70 80 90 100 110 120

gi-50913505. QTETQGTQVVATKQKLESENSSLKVALKRESGFEHNATIDASLDTE
70 80 90 100 110 120

gi-50913505. VTMALLELRKQGLSIVDTKIVRIQSSTNQRNDITTLTFKNGLSLEG
130 140 150 160 170 180

gi-50913505. VTMALLELRKQGLSIVDTKIVRIQSSTNQRNDITTLTFKNGLSLEG
130 140 150 160 170 180

gi-50913505. VNPNDTVQTITPTIKQDADGKVNLVFTGRLGKQVIIIVSTTRLKEE
190 200 210 220 230 240

gi-50913505. VNPNDTVQTITPTIKQDADGKVNLVFTGRLGKQVIIIVSTTRLKEE
190 200 210 220 230 240

gi-50913505. GAVGLSQDKDRPPYSKPITVNILKPKLSSIESLDSKDFEIVKTID
250 260 270 280 290 300

gi-50913505. GAVGLSQDKDRPPYSKPITVNILKPKLSSIESLDSKDFEIVKTID
250 260 270 280 290 300

gi-50913505. SKQYEVLKTDYQSAKDSTPQTRDILFGEYTVPLVMNKGHNN
310 320 330 340 350 360

gi-50913505. SKQYEVLKTDYQSAKDSTPQTRDILFGEYTVPLVMNKGHNN
310 320 330 340 350 360

gi-50913505. AAPALIQPRSFRSLTPRSTRMKRSAPVEKFEGELEHHKRIDL
370 380 390 400 410 420

gi-50913505. AAPALIQPRSFRSLTPRSTRMKRSAPVEKFEGELEHHKRIDL
370 380 390 400 410 420

gi-50913505. HDTSDLYRLYLDMTGKKNPLDILVVVDKSGSMQE
430 440 450 460 470 480

gi-50913505. HDTSDLYRLYLDMTGKKNPLDILVVVDKSGSMQE
430 440 450 460 470 480

gi-50913505. TFDYSSYQGESFNRCQIHYRYRGIVSVSDGIRRDAVKNSLLGV
490 500 510 520 530 540

gi-50913505. TFDYSSYQGESFNRCQIHYRYRGIVSVSDGIRRDAVKNSLLGV
490 500 510 520 530 540

gi-50913505. SVIGFQGSADYHAGKWP
550 560 570 580 590 600

gi-50913505. SVIGFQGSADYHAGKWP
550 560 570 580 590 600

gi-50913505. GTNYHAALLKAKEILNEVKDDGRRKIMIF
610 620 630 640 650 660

gi-50913505. GTNYHAALLKAKEILNEVKDDGRRKIMIF
610 620 630 640 650 660

FIGURE 57B

PCT/US2005/027239

gi-50913505. QEGSKLAIDEFKARYPNLSIYSLGVSKDINSDTASSSPVVLKYLSEEHYYGITDAALE
 ||||| ||||| ||||| ||||| ||||| |||||
 gi-50913505. QEGSKLAIDEFKARYPNLSIYSLGVSKDINSDTASSSPVVLKYLSEEHYYGITDAALE
 670 680 690 700 710 720

gi-50913505. KTLNKIVEDSKLSQLGISDSLSQYVDYYDKQPDVLVTRKSKVNDETEILYQKDQVQEAGK
 ||||| ||||| ||||| ||||| ||||| |||||
 gi-50913505. KTLNKIVEDSKLSQLGISDSLSQYVDYYDKQPDVLVTRKSKVNDETEILYQKDQVQEAGK
 730 740 750 760 770 780

gi-50913505. DIIDKVVFTPCKTTSQPKGVTLTFKSDYKVDDETYTLISFNVKASDEAYEKYKDNEGRYS
 ||||| ||||| ||||| ||||| |||||
 gi-50913505. DIIDKVVFTPCKTTSQPKGVTLTFKSDYKVDDETYTLISFNVKASDEAYEKYKDNEGRYS
 790 800 810 820 830 840

gi-50913505. EMGDSDTDYGTNQTSSGKGGLPSNSDASVNYMADGREQKLKYKHPVIQVKTVPIFTKVD
 ||||| ||||| ||||| ||||| |||||
 gi-50913505. EMGDSDTDYGTNQTSSGKGGLPSNSDASVNYMADGREQKLKYKHPVIQVKTVPIFTKVD
 850 860 870 880 890 900

gi-50913505. ADNNQKKLAGVEFELRKEDKKIVWEKGTTGSNGQLNFKYLQKGKTYLYETKAKLGYTLP
 ||||| ||||| ||||| ||||| |||||
 gi-50913505. ADNNQKKLAGVEFELRKEDKKIVWEKGTTGSNGQLNFKYLQKGKTYLYETKAKLGYTLP
 910 920 930 940 950 960

gi-50913505. ENPWEVAVANNGDIKVKHPIEGELKSKDGSYMIKNYKIYQLPSSGGRGSQIFIIIVGSMTA
 ||||| ||||| ||||| |||||
 gi-50913505. ENPWEVAVANNGDIKVKHPIEGELKSKDGSYMIKNYKIYQLPSSGGRGSQIFIIIVGSMTA
 970 980 990 1000 1010 1020

gi-50913505. TVALLFYRRQHRKKQY
 ||||| |||||
 gi-50913505. TVALLFYRRQHRKKQY
 1030

gi-50913505.pep
 /home/morama/gas/pili/align/gi-19224141.pep

gi|19224141|gb|AAL86412.1|AF447492_9 protein F2 [Streptococcus pyogenes]

SCORES : Init1: 63 Initn: 100 Opt: 159 z-score: 105.9 E(): 0.023
 >>/home/morama/gas/pili/align/gi-19224141.pep (1161 aa)
 initn: 100 init1: 63 opt: 159 z-score: 105.9 expect(): 0.023
 Smith-Waterman score: 159; 36.7% identity in 98 aa overlap
 (895-990:48-144)

gi-50913505. SDASVNYMADGREQKLKYKHPVIQVKTVPIFTKVDADNNQKKLAGVEFELRKEDKK-IV
 :| :|:|:| | :| | :| :|
 gi-19224141. FILGLLLFIGLGSVSVGHAETRNGANKQGSFEIKVWDQNNKPLPGATFSLTSKDGKGT
 20 30 40 50 60 70

930 940 950 960 970 980

FIGURE 57C

PC-T/US005/Z7339

gi-50913505. WEKGTTGSNGQLNFYLYQKGTYYLYETKAKLGYTLPEVAVANNGDIK-VKHPIEG
 : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 gi-19224141. VQTFTSNDKGIVDAQNLQPG-TYTLKEETAPDGYDKTSRTWTVIVYENGYTKLVENPYNG
 80 90 100 110 120 130

990 1000 1010 1020 1030
 gi-50913505. ELKSKDGSYMIKNYKTYQLPSSGGRGSQIFIIIVGSMTATVALLFYRRQHRKKQY
 | : | | : |
 gi-19224141. EIISKAGSKDVSSLQLENPKMSVVSKEYGKTEVSSGAADFYRNHAAYFKMSFELKQKDKS
 140 150 160 170 180 190

gi-50913505.pep

/home/morrama/gas/pili/align/gi-21909640.pep

gi|21909640|ref|NP_663908.1| protein F2 like fibronectin-binding protein [Streptococcus pyogenes MGAS315]

SCORES Init1: 35 InitN: 35 Opt: 136 z-score: 96.3 E(): 0.08
 >>/home/morrama/gas/pili/align/gi-21909640.pep (656 aa)
 initN: 35 init1: 35 opt: 136 z-score: 96.3 expect(): 0.08
 Smith-Waterman score: 148; 24.5% identity in 339 aa overlap
 (686-1005:147-449)

660 670 680 690 700 710
 gi-50913505. NVTRSQEGLKLAIDEFKARYPNLSTIYSLGVSKDINSDTASSSPVVL----KYLSGEEHYY
 : | | | : : : | : : : | : | : | : | : | : | : | : | : | : | : | : |
 gi-21909640. TNLQAVISVEPVIESLPWTSLKPIAQKDITAKKIWVDAPKEKPIIYFKLYRQLPGEKEV-
 120 130 140 150 160 170

720 730 740 750 760 770
 gi-50913505. GITDTAELEKTLNKIVEDSKLSQLGISDLSQYVVDYYDKQPDVLTFRSKVNDTEILYQ
 : | | | : : : : : : | : | : | : | : | : | : | : | : | : | : | : |
 gi-21909640. -AVDDAELQ----INSEGQQEISVTWT-NQLVT--DEKGMAITYSVKEVDKNGELLEP
 180 190 200 210 220

780 790 800 810 820 830
 gi-50913505. KDQVQ-EAGKDIIDKVVFTPCKTTSQPKGVTLTFKSDYKVDEYTYTLSFNVKASDEAYE
 | | : | | : : | | | : : | : | : | : | : | : | : | : | : | : | : | : |
 gi-21909640. KDYIKKEDGLTVNTYV--KPTSG-HYDIEVTFGNGH-ID--TEDTTTPDIVSGEN---
 230 240 250 260 270

840 850 860 870
 gi-50913505. KYKDNEGRYSEMGSSTDYGTNQTSSGKGGLPSNSDA-SVNY-MADGR-----
 : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 gi-21909640. QMKQIEGEDSKPIDEVTE--NNLIEFGKNTMPGEEDGTNSNKYEVEDSRPVDTLSGLSS
 280 290 300 310 320 330

880 890 900 910 920 930
 gi-50913505. EQKLKYKHPVIQVKTPITFTKVDADNNQQKLAGVEFLRKEDKKIVWEKGTTGSNGQIN
 | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 gi-21909640. EQGQSGDMTIEEDSATHLKF SKRDTDG--KELAGATMELRDSSGKTI--STWISDGQVK
 340 350 360 370 380

940 950 960 970 980 990
 gi-50913505. FKYLQKGKTYLYETKAKLGYTLPEVAVANNGDIKVKHPIEGELKSKDGSYMIKNY
 | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 gi-21909640. DFYLMMPGK-YTFVETAAPDGY-----EVATAITFTVNEQQVTVNGKATKGDAAHIVMV
 390 400 410 420 430 440

1000 1010 1020 1030
 gi-50913505. KIYQLPSSGGRGSQIFIIIVGSMTATVALLFYRRQHRKKQY

FIGURE 57D

gi-21909640. DAYK-PTKGSGQVIDIEEKLPEQGHSGSTETEDSKSSDVIIIGGQGEVVDTTEDTQSGM
450 460 470 480 490 500

gi-50913505.pep

/home/morama/gas/pili/align/gi-13621428.pep

gi|13621428|gb|AAK33238.1| hypothetical protein [Streptococcus pyogenes]

```
SCORES Init1: 33 Initn: 33 Opt: 91 z-score: 75.6 E(): 1.1  
>/home/morama/gas/pili/align/gi-13621428.pep (340 aa)  
initn: 33 init1: 33 opt: 91 z-score: 75.6 expect(): 1.1  
Smith-Waterman score: 95; 19.9% identity in 271 aa overlap  
(568-819:57-318)
```

540	550	560	570	580	590			
gi-50913505.	NKLSVIGFQGSADYHAGKWPDQSPRGFYQPQNLNNNSRDAELLKGWSTNSLLDPNTLTAL							
			: : : : : : : : : :					
gi-13621428.	VNGAKLTVTKNLSDLVNSNALIPMTDFTFKIEPDPTVNEDGNKFKGVALNTPMTKVTYTMS							
			30	40	50	60	70	80

	660	670	680	690	700	
gi-50913505.	SNDRNNVTRSQE GSKLAI DEFKARYPNL SIYSLGVSKD INSDTAS-SSPVVLKVYL SGEEH	:	:	:		
gi-13621428.	KPVATYIVGYKE GSKVPI -QFKN---S LDSTTLTVKKVSGTGGDRSKDFNFGLTLKANQ	:	:	:		
	150	160	170	180	190	200

710	720	730	740	750		
gi-50913505.	YYGITDTAELEKTLN---	KIVEDSKLSQL---	GISDSDLSQLYVDYYDKQPDVLVT	--R		
	:: : : . : . : ; : : : . : : :					
gi-13621428.	YYKASEKVMIIEKTTKGQAPVQTEASIDQLYHFTLKDGESIKVTNLPGVGDYVVTEDDYK					
	210	220	230	240	250	260

	760	770	780	790	800	810
gi-50913505	KSKVNDETEIILYQKDQVQEAGKDEIIDKVVFPTKTTSQPKGVTLTFKSDYKVVDDEYTYTL	:	:	:	:	:
gi-13621428	SEKYTTNVEVSPQDGAVRNIAQMSSTEQETSTDKDMT	—	—	—	TTFTNKKDFEVPTGVAMTV	
	270	280	290	300	310	

820 830 840 850 860 870
gi-50913505. SFNVKASDEAYEKYKDNEGRYSEMGDSDTDYGTNQTSSGKGGLPSNSDASVNYMADGREQ
:
gi-13621428. APYTALGIVAVGGALYFVKKKNA
320 330 340

gi-50913505.pep

/home/morama/gas/pili/align/gi-50913506.pep

gi|50913506|ref|YP_059478.1| Fimbrial structural subunit [Streptococcus pyogenes MGAS10394].

SCORES Init1: 70 Initn: 149 Opt: 86 z-score: 71.3 E(): 1.9
>>/home/morama/gas/pili/align/gi-50913506.pep (556 aa)

FIGURE 57E

initn: 149 initl: 70 opt: 86 z-score: 71.3 expect(): 1.9
 Smith-Waterman score: 120; 21.5% identity in 469 aa overlap
 (503-966:33-428)

480	490	500	510	520	
gi-50913505. YSQWVYHGTFDYSSYQGESFNRGQIHYRYRGIVSVSDGIRRDAAVKNSLLGVNG-----L					
:: :: : :: : ::					
gi-50913506. NRRETVREKILITAKKMLACLAIALAVVGLGMTRVS-ALSKDDTAQLKITNIEGGPTVTL					
10	20	30	40	50	60
530	540	550	560	570	580
gi-50913505. LQRFVNINPENKLSVIGFQGSADYHAGKWPQSPRGFYQPMLNNSRDAELLKGWSTNS					
: : : : : : : : : : : :					
gi-50913506. YKIGEGVYNTNGDSFINFK---YAEGVSLTETGPTSQEIT-TIANGINTGKIKPFSSEN					
70	80	90	100	110	
590	600	610	620	630	640
gi-50913505. LLDPNTLTALHNNGTNYHAALLKAKEILNEVKDDGRRKIMIFTISDGVPFTYFGEDGYRSG					
: :: : :: : : : : : : : :					
gi-50913506. VSISNGTATYNARGASVYIALLTGAT-----DGRTYNPILLAAS----YNGEGNLVTK					
120	130	140	150	160	
650	660	670	680	690	700
gi-50913505. NGSSNDRNNVTRSQECKLAIDEFKARYPNLSTYSLGVSKDINSDTASSSPVVLKYLSGE					
: :: : :: : : : : : : :					
gi-50913506. NIDS--KSNLYLGQ--TSVA---KSSLPSITKKVTGTTIDDVNKKTTSLGSVLSYSLTFE					
170	180	190	200	210	
710	720	730	740	750	760
gi-50913505. EHYYGITDTAELEKTLNKIVEDSKLSQLGIISDLSQYVDYDKQPDVLVTRKSKVNDTE					
: :: : : : : : : : : :					
gi-50913506. LPSY-----TKEAVNKTVY-----VSDNMSEGLTF--NFNSLTVEWKGMANITE					
220	230	240	250	260	
770	780	790	800	810	820
gi-50913505. ILYQKDQVQEAGKDIIDKVVFPTKTTSQPKGKVTLTFSYKVDDEYTYSFNVKASDE					
:: : : : : : : :					
gi-50913506. -----DCSVMVENTKIGIAKEVNNGFNLSFIYDS--LESISPNI-----					
270	280	290			
830	840	850	860	870	880
gi-50913505. AYEKYKDNEGRYSEMGDSDTDYGTNQTSSGKGGLPSNSDASVNYMADGREQKLPYKHPVI					
: : :: : : : : : : : :					
gi-50913506. SYKAVVNINKATVCECNPNAEFFYSNNPTKGNTYDNLDDKKPDK-GNGITSKEDSK					
300	310	320	330	340	350
890	900	910	920	930	940
gi-50913505. QVKTVPLTFKVDADNNQKKLAGVEFELRKEDKKIVWEKGTTGNGQLNFKYLOKGKTYY					
: : : : : : : : : :					
gi-50913506. IVYTYQIAFRKVDS-VSKTPLIGAIFGVYDTSNKL-LDIVTINKNGYAITSTQVSSGK-YK					
360	370	380	390	400	
950	960	970	980	990	1000
gi-50913505. LYETKAALKGYTLPENPWEVAVANNGDIKVKHPIEGELKSKDGSYMIKNYKITYQLPSSGR					
: : : : :					
gi-50913506. IKELKAPKGYSLNTETYEITANWVTATVKTSAKSSTTYTSDKNKATDNSEQVGWLKNGI					
410	420	430	440	450	460

gi-50913505.pep

/home/morama/gas/pili/align/gi-13621432.pep

gi|13621432|gb|AAK33241.1| conserved hypothetical protein [Streptococcus pyogene]

FIGURE 57F

PCT/US2005/027239 181/487

s]

SCORES Init1: 40 Initn: 65 Opt: 78 z-score: 68.0 E(): 2.9
 >>/home/morama/gas/pili/align/gi-13621432.pep (450 aa)
 initn: 65 init1: 40 opt: 78 z-score: 68.0 expect(): 2.9
 Smith-Waterman score: 78; 37.0% identity in 46 aa overlap
 (368-411:14-56)

	340	350	360	370	380	390
gi-50913505.	KGHNNTINIIYIRSTRPLGLKPIGAAPALIQPRSFRSLTPRSTRMK--RSAPVEKFEGELE					
gi-13621432.	MTRTNYQKKRMTCPVETEDITYRRKKIKGRRQAILAQFEPELV					
	10	20	30	40		

	400	410	420	430	440	450
gi-50913505.	HHKRIDYLGDNQNPNPTTIDDKDEHDTSDLYRLYLDLMTGKKNPLDIEVVVDKSGSMQEG					
gi-13621432.	HHELI---GDSCTCPDCHGTLTEIGSVVQRQELVFIAPAQLKRINHVQHAYKCQTCSDNSL					
	50	60	70	80	90	100

gi-50913505.pep

/home/morama/gas/pili/align/gi-19745301.pep

gi|19745301|ref|NP_606437.1| putative collagen binding protein [Streptococcus pyogenes MGAS8232]

SCORES Init1: 52 Initn: 52 Opt: 73 z-score: 64.8 E(): 4.3
 >>/home/morama/gas/pili/align/gi-19745301.pep (524 aa)
 initn: 52 init1: 52 opt: 73 z-score: 64.8 expect(): 4.3
 Smith-Waterman score: 95; 23.7% identity in 245 aa overlap
 (759-989:241-466)

	730	740	750	760	770	780
gi-50913505.	DSKLSQLGISDSLSQLYDVKQPDVLVTRKSKVNDETEILYQKDQVQEAG-KDIIDKVV					
gi-19745301.	ETIDPDFNEGKEIKYTHILGADLFYSANNPRASTNDE--LLSQVKVLEKGYRD--DSTT					
	220	230	240	250	260	
	790	800	810	820	830	840
gi-50913505.	FTPKTTSQPKGVTLT--FKSDYKVDD-EYTYTLSFNVKASDEAYEKYKDNEYRYSEM					
gi-19745301.	YANLTSVEFRAATOLAIYYFTDSVLDNLADYHGFALTTEALNATKETVAYAEDRANLP					
	270	280	290	300	310	320
	850	860	870	880	890	
gi-50913505.	GDSDTDY--GTNOTSS-GKGGLPSNSDAVNMYADGREQKLKYKHPVIQVKTVPTIFT					
gi-19745301.	NISNLDFYVPNSNKYQSLIGTOYHP-ESLVDLIERMEDKQAPIIPTITHKLTIISKTVTGTI-					
	330	340	350	360	370	380
	900	910	920	930	940	950
gi-50913505.	KVDADNNQQKKLAGVFELRKEDKKIVWEKGTTGSN-GQLNFKYLQKGK-TYYLYETKARL					
gi-19745301.	---AD--KKKEFNFEIHLKSSDGQAI--SGTYPTNSGELT---VTDGKATFTLKDGESLI					
	390	400	410	420	430	
	960	970	980	990	1000	1010
gi-50913505.	GYTLPEN-PWEVAVANNGDIKVHKPIEGELKSKDGSYMIKNYKIYQLPSSGGRGSQIFII					
gi-19745301.	VEGLPSGYSYEITETGASDYEV--VNGK-NAPDGKATKASVKEDETITFENRKDLVPPT					

FIGURE 57G

PCT/US05/27239

440 450 460 470 480 490

1020 1030

gi-50913505. VGSMTATVALLFYRRQHRKKQY

gi-19745301. GLTTDGAIYLWLLLIVLLGLWWLIGRKGLKND
500 510 520

gi-50913505.pep

/home/morama/gas/pili/align/gas15.pep

GAS15 GAS15

SCORES Init1: 43 Initn: 68 Opt: 69 z-score: 61.4 E(): 6.6
 >>/home/morama/gas/pili/align/gas15.pep (762 aa)
 initn: 68 init1: 43 opt: 69 Z-score: 61.4 expect(): 6.6
 Smith-Waterman score: 100; 21.4% identity in 252 aa overlap
 (641-873:492-739)

620 630 640 650 660
 gi-50913505. AKEILNEVKDDGRRKIMIFISDGVPFTFYFGEDGYRSGNGSSNDRNNVTRSQ--EGSKLAI
 |||| :|:: : ::::|::| ::::||
 gas15.pep HTAGRDLFKYTVKPRDTDPDTFLKHKKVIEKGYRE-KGQALEYESGLTETQLRAATQLAI
 470 480 490 500 510 520

670 680 690 700 710 720
 gi-50913505. DEF--KARYPNLSIYSLGVSKDINSDTASSSPVVLKYLSEEHYYGITDTAELEKTLNKI
 ||| :|: : :: : |::| : : ::| :| :| :| :| :|
 gas15.pep YYFTDSAELDKDLKDYHGFGDMNDSTLAVAKILVEY-AQDSNPQQLTDLDFFIPNNNKY
 530 540 550 560 570

730 740 750 760 770
 gi-50913505. VEDSKLSQLGISDSLSQLYVDDYDKQPDVLVT----RKSKVN---DETETLYQKDQVQEAE
 :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
 gas15.pep --QSLIGTQWHPEDLVDIIRMEDKKEVIPVTHNLTLRKTVTGLAGDRTKDFHFEIELKNN
 580 590 600 610 620 630

780 790 800 810 820 830
 gi-50913505. GKDIIIDKVVFTPKTTISQPK-GKVTLTFKSDYKVDDE-YTYTLSFNVKASDEAYEKYKDNE
 :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
 gas15.pep KQELLSOTVKTDKTNLEFKDGKATINLKHEGESLTQQLPEGYSYLVKETDSEGYKVVKVNS
 640 650 660 670 680 690

840 850 860 870 880 890
 gi-50913505. GRYSEMGSDDTDYGTNQT---SSGKGGLPSNSDASVN-YMADGREQKLPHYKHPVIQVKT
 :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
 gas15.pep QEVANATVSKTGITSDETAFENNKEPVVPTGVDQKINGYLALIVIAGISLGIWGIHTIR
 700 710 720 730 740 750

900 910 920 930 940 950
 gi-50913505. VPITFTKVDADNNQKIKLAGVEFELRKEDKKIVWEKGTTGSNGQLNFKYLQKGKTYLYET
 gas15.pep IRKHD
 760

gi-50913505.pep

/home/morama/gas/pili/align/gi-21909636.pep

gi|21909636|ref|NP_663904.1| conserved hypothetical protein [Streptococcus pyogenes MGAS315]

SCORES Init1: 31 Initn: 31 Opt: 62 z-score: 60.8 E(): 7.1
 >>/home/morama/gas/pili/align/gi-21909636.pep (344 aa)
 initn: 31 init1: 31 opt: 62 z-score: 60.8 expect(): 7.1
 Smith-Waterman score: 71; 22.9% identity in 131 aa overlap
 (181-305:176-298)

	160	170	180	190	200	210
gi-50913505.	DITTTLTFKNGLSLEGASTEANDPNVRVGIVNPNDTVQTITPTIKQDADGKVKNLVFTGR					
	150	160	170	180	190	
gi-21909636.	VDVYVGNKEGGGFEPKFIIVSKEQGTDVKKPVNFMNSFATTSLVKKNVSGN-----TGE					
	220	230	240	250	260	
gi-50913505.	LGKQ----VIIVSTTRLKEEQTISLD SYGELVIDGAVGLSQKDRPPYSKPITVNILKPKL					
	200	210	220	230	240	250
gi-21909636.	LQKEFDFTTLNESTNFKKDQIVSLQK-GNEKFEVKIGTPYFKFLKGESIQLDKLPVGI					
	270	280	290	300	310	320
gi-50913505.	SSIESSLDSKDFEIVKTIDNLYTWDDQ--FYLLDFISKQYEVLKTDYQSAKDSTPQTRDI					
	260	270	280	290	300	310
gi-21909636.	TYKVNEMEANK-DGYKTTASLKEGDGQSKMYQLDMEQKTDESADEIVVTNKRDTQVPTGV					
	330	340	350	360	370	380
gi-50913505.	LFGEYTVEPLVMNKGHNNNTINIIYIRSTRPLGLKPIGAAPALIQPRSFRSLTPRSTRMKRS					
gi-21909636.	VGTLAPFAVL SIVAIGGVIYITKRKKA					
	320	330	340			

gi-50913505.pep

/home/morama/gas/pili/align/gi-28810259.pep

gi|28810259|dbj|BAC63197.1| hypothetical protein [Streptococcus pyogenes SSI-1]

SCORES Init1: 31 Initn: 31 Opt: 62 z-score: 60.7 E(): 7.2
 >>/home/morama/gas/pili/align/gi-28810259.pep (349 aa)
 initn: 31 init1: 31 opt: 62 z-score: 60.7 expect(): 7.2
 Smith-Waterman score: 71; 22.9% identity in 131 aa overlap
 (181-305:181-303)

	160	170	180	190	200	210
gi-50913505.	DITTTLTFKNGLSLEGASTEANDPNVRVGIVNPNDTVQTITPTIKQDADGKVKNLVFTGR					
	160	170	180	190	200	
gi-28810259.	VDVYVGNKEGGGFEPKFIIVSKEQGTDVKKPVNFMNSFATTSLVKKNVSGN-----TGE					
	220	230	240	250	260	
gi-50913505.	LGKQ----VIIVSTTRLKEEQTISLD SYGELVIDGAVGLSQKDRPPYSKPITVNILKPKL					
	210	220	230	240	250	260
gi-28810259.	LQKEFDFTTLNESTNFKKDQIVSLQK-GNEKFEVKIGTPYFKFLKGESIQLDKLPVGI					
	270	280	290	300	310	320
gi-50913505.	SSIESSLDSKDFEIVKTIDNLYTWDDQ--FYLLDFISKQYEVLKTDYQSAKDSTPQTRDI					
	270	280	290	300	310	320
gi-28810259.	TYKVNEMEANK-DGYKTTASLKEGDGQSKMYQLDMEQKTDESADEIVVTNKRDTQVPTGV					

FIGURE 57I

PCT/US05/237239

330 340 350 360 370 380

gi-50913505. LFGEYTVEPLVMNKGNNTINIIYIRSTRPLGLKPIGAAPALIQPRSFRSLTPRSTRMKRS

gi-28810259. VGTLAPFAVLSIVAIIGGVIYITKRKKA

330 340

gi-50913505.pep

/home/morama/gas/pili/align/gi-19224139.pep

gi|19224139|gb|AAL86410.1|AF447492_7 unknown [Streptococcus pyogenes]

SCORES Init1: 43 Initn: 43 Opt: 54 z-score: 58.9 E(): 8.9
 >>/home/morama/gas/pili/align/gi-19224139.pep (189 aa)

initn: 43 init1: 43 opt: 54 z-score: 58.9 expect(): 8.9

Smith-Waterman score: 54; 31.6% identity in 57 aa overlap
 (742-796:90-143)

720 730 740 750 760

gi-50913505. GITDTAELEKTLNKIVEDSKLSQLGISDLSQYVVDYYDKQP--DVLVTRKSKVNDETEIL

|| || || : || |

gi-19224139. ITIAGSGKASFSPLTFTTVGQTYRKYQKPSQNKDYQADTTVFDVLVYV--TYDEDGTL

60 70 80 90 100 110

770 780 790 800 810 820

gi-50913505. YQKDQVQEAGKDIIDKVVFPTPKTTSQPKGVTLTFKSDYKVDDETYTLFSNVKASDEAY

| : || : : || | : |

gi-19224139. VAKVISRAGDEEKSAITFKPKRLVKPIPQPNIPKTPPLPLAGEVKSLGILSIVLLGL

120 130 140 150 160 170

gi-50913505.pep

/home/morama/gas/pili/align/gi-19745305.pep

gi|19745305|ref|NP_606441.1| hypothetical protein [Streptococcus pyogenes MGAS82 32]

SCORES Init1: 43 Initn: 43 Opt: 54 z-score: 58.8 E(): 9
 >>/home/morama/gas/pili/align/gi-19745305.pep (195 aa)

initn: 43 init1: 43 opt: 54 z-score: 58.8 expect(): 9

Smith-Waterman score: 54; 31.6% identity in 57 aa overlap
 (742-796:96-149)

720 730 740 750 760

gi-50913505. GITDTAELEKTLNKIVEDSKLSQLGISDLSQYVVDYYDKQP--DVLVTRKSKVNDETEIL

|| || || : || |

gi-19745305. ITIAGSGKASFSPLTFTTVGQTYRKYQKPSQNKDYQADTTVFDVLVYV--TYDEDGTL

70 80 90 100 110 120

770 780 790 800 810 820

gi-50913505. YQKDQVQEAGKDIIDKVVFPTPKTTSQPKGVTLTFKSDYKVDDETYTLFSNVKASDEAY

| : || : : || | : |

gi-19745305. VAKVISRAGDEEKSAITFKPKRLVKPIPQPDIPKTPPLPLAGEVKSLGILSIVLLGL

130 140 150 160 170 180

gi-50913505.pep

/home/morama/gas/pili/align/orf82.pep

TRANSLATE of: orf82.seq check: 4296 from: 1 to: 672
 generated symbols 1 to: 224.

PCT/US05/237239
GETSEQ from morama, September 13, 2004 17:09.

SCORES Init1: 43 Initn: 43 Opt: 54 z-score: 58.2 E(): 9.6
>>/home/morama/gas/pili/align/orf82.pep (224 aa)
initn: 43 init1: 43 opt: 54 Z-score: 58.2 expect(): 9.6
Smith-Waterman score: 54; 31.6% identity in 57 aa overlap
(742-796:125-178)

	720	730	740	750	760	
gi-50913505.	GITDTAELEKTLNKIVEDSKLSQLGISDSLSQYVVDYYDKQP--	DVLVTRKSKVNDTEIL				
orf82.pep	ITIAGSGKASFPLTFTTVGQTYTYRKYQKPSQNQDYQADTTVFDVLVYV--	TYDEDGTL				
	100	110	120	130	140	
					150	
	770	780	790	800	810	820
gi-50913505.	YQKDQVQEAGKDIIDKVVFTPCKTSQPKGVTLTFKSDYKVDDETYTLFSFNVKASDEAY					
orf82.pep	VAKVISRRAGDEEKSATTFKPKRLVVKPIPPRQPNIPKTPPLPLAGEVKSLLGILSIVLLGL					
	160	170	180	190	200	210

gi-50913505.pep
/home/morama/gas/pili/align/gi-21909638.pep

gi|21909638|ref|NP_663906.1| hypothetical protein [Streptococcus pyogenes MGAS315]

SCORES Init1: 43 Initn: 43 Opt: 52 z-score: 58.0 E(): 9.9
>>/home/morama/gas/pili/align/gi-21909638.pep (183 aa)
initn: 43 init1: 43 opt: 52 Z-score: 58.0 expect(): 9.9
Smith-Waterman score: 52; 31.6% identity in 57 aa overlap
(742-796:84-137)

	720	730	740	750	760	
gi-50913505.	GITDTAELEKTLNKIVEDSKLSQLGISDSLSQYVVDYYDKQP--	DVLVTRKSKVNDTEIL				
gi-21909638.	ITIAGSGKASFPLTFTTVGQTYTYRKYQKPSQNQDYQADTTVFDVLVYV--	TYDEDGTL				
	60	70	80	90	100	
					110	
	770	780	790	800	810	820
gi-50913505.	YQKDQVQEAGKDIIDKVVFTPCKTSQPKGVTLTFKSDYKVDDETYTLFSFNVKASDEAY					
gi-21909638.	VAKVISRRAGDEEKSATTFKPKRLVVKPIPPRQPNIPKTPPLPLAGEVKSLLGILSIVLLGL					
	120	130	140	150	160	170

! Distributed over 1 thread.
! Start time: Wed Sep 15 18:46:18 2004
! Completion time: Wed Sep 15 18:46:21 2004

! CPU time used:
Database scan: 0:00:00.1
Post-scan processing: 0:00:02.2
Total CPU time: 0:00:02.3
Output File: gi-50913505.fasta

PCT/US05/27239
SEQUENCE LIST 1.0

(Peptide) FASTA of: gi-50913506.pep from: 1 to: 556 September 15, 2004 18:45

gi|50913506|ref|YP_059478.1| Fimbrial structural subunit [Streptococcus pyogenes MGAS10394]

TO: *.pep Sequences: 56 Symbols: 22,803 Word Size: 2

Scoring matrix: GenRunData:blosum50.cmp

Variable pamfactor used

Gap creation penalty: 12 Gap extension penalty: 2

Histogram Key:Each histogram symbol represents 1 search set sequences
z-scores computed from opt scores

z-score	obs	exp
(=)	(*)	
< 20	0	0:
22	0	0:
24	0	0:
26	0	0:
28	0	0:
30	0	0:
32	0	0:
34	0	1:*
36	0	2: *
38	0	3: *
40	0	4: *
42	0	5: *
44	0	5: *
46	1	5:= *
48	8	5:====*==
50	24	5:====*=====
52	2	4:== *
54	3	3:==*
56	1	3:= *
58	2	2:=*
60	3	2:=*=
62	0	2:= *
64	2	1:=*=
66	2	1:=*=
68	1	1:=*
70	2	1:=*=
72	0	0:
74	3	0:====
76	1	0:=
78	0	0:
80	0	0:
82	0	0:
84	0	0:
86	0	0:
88	0	0:
90	0	0:
92	0	0:
94	0	0:
96	0	0:
98	0	0:

FIGURE 58

100	0	0:
102	0	0:
104	0	0:
106	0	0:
108	0	0:
110	0	0:
112	0	0:
114	0	0:
116	0	0:
118	0	0:
>120	1	0:=

Joining threshold: 37, opt. threshold: 25, opt. width: 16, reg.-scaled

The best scores are:

		init1	initn	opt	z-sc E(55)
/home/morrama/gas/pili/align/gi-50913506.pep		Begin: 1	End: 556		
! gi 50913506 ref YP_059478.1 Fimbri...	3454	3454	3454	1016.2	4.7e-53
/home/morrama/gas/pili/align/orf84.pep		Begin: 316	End: 567		
! TRANSLATE of: orf84.seq check: 7868...	57	83	135	75.1	1.2
/home/morrama/gas/pili/align/gi-19745307.pep		Begin: 316	End: 567		
! gi 19745307 ref NP_606443.1 protei...	57	83	135	75.1	1.2
/home/morrama/gas/pili/align/gi-21909640.pep		Begin: 202	End: 524		
! gi 21909640 ref NP_663908.1 protei...	56	81	134	75.0	1.2
/home/morrama/gas/pili/align/gi-28810263.pep		Begin: 316	End: 638		
! gi 28810263 dbj BAC63201.1 protein...	56	82	134	74.7	1.3
/home/morrama/gas/pili/align/orf80.pep		Begin: 49	End: 352		
! TRANSLATE of: orf80.seq check: 9824...	45	69	113	70.8	2.1
/home/morrama/gas/pili/align/gi-19224137.pep		Begin: 25	End: 342		
! gi 19224137 gb AAL86408.1 AF447492...	45	69	109	69.8	2.4
/home/morrama/gas/pili/align/gi-19224141.pep		Begin: 277	End: 645		
! gi 19224141 gb AAL86412.1 AF447492...	73	73	118	68.9	2.7
/home/morrama/gas/pili/align/gi-21909636.pep		Begin: 44	End: 344		
! gi 21909636 ref NP_663904.1 cons...	45	98	96	66.1	3.8
/home/morrama/gas/pili/align/gi-28810259.pep		Begin: 49	End: 349		
! gi 28810259 dbj BAC63197.1 hypothe...	45	98	96	66.0	3.8
/home/morrama/gas/pili/align/gas15.pep		Begin: 222	End: 470		
! GAS15_GAS15		42	68	96	63.8
					5
/home/morrama/gas/pili/align/gi-13621428.pep		Begin: 17	End: 340		
! gi 13621428 gb AAK33238.1 hypothet...	41	41	87	63.6	5.2
/home/morrama/gas/pili/align/gi-19224135.pep		Begin: 193	End: 462		
! gi 19224135 gb AAL86406.1 AF447492...	41	41	86	61.0	7
/home/morrama/gas/pili/align/gi-50913505.pep		Begin: 503	End: 966		
! gi 50913505 ref YP_059477.1 Collag...	70	149	86	60.1	7.8
/home/morrama/gas/pili/align/gi-13621430.pep		Begin: 60	End: 143		
! gi 13621430 gb AAK33240.1 hypothet...	43	67	67	59.2	8.7
/home/morrama/gas/pili/align/gi-19745303.pep		Begin: 44	End: 344		
! gi 19745303 ref NP_606439.1 hypoth...	51	106	69	58.4	9.5
\End of List					

gi-50913506.pep

/home/morrama/gas/pili/align/gi-50913506.pep

gi|50913506|ref|YP_059478.1| Fimbrial structural subunit [Streptococcus pyogenes MGAS10394]

SCORES Init1: 3454 Initn: 3454 Opt: 3454 z-score: 1016.2 E(): 4.7e-53
 >>/home/morrama/gas/pili/align/gi-50913506.pep (556 aa)
 initn: 3454 init1: 3454 opt: 3454 Z-score: 1016.2 expect(): 4.7e-53
 Smith-Waterman score: 3454; 100.0% identity in 556 aa overlap

FIGURE 58A

PCT/US05/27239
(1-556:1-556)

	10	20	30	40	50	60
gi-50913506.	MTNRRET	VREKILITAKK	IMLACLA	LAVVGLGMTR	VSALS	KDDTAQLKITNIEGGPTVT
gi-50913506.	MTNRRET	VREKILITAKK	IMLACLA	LAVVGLGMTR	VSALS	KDDTAQLKITNIEGGPTVT
	10	20	30	40	50	60
	70	80	90	100	110	120
gi-50913506.	LYKIGEGVYNTNGDS	FINFKYAEGVSL	TETGPTSQEIT	TIANGINTGKIKPF	STENVSIS	
gi-50913506.	LYKIGEGVYNTNGDS	FINFKYAEGVSL	TETGPTSQEIT	TIANGINTGKIKPF	STENVSIS	
	70	80	90	100	110	120
	130	140	150	160	170	180
gi-50913506.	NGTATYNARGASVYI	ALLTGATDGR	TYNPILLAASYN	GEGLVTKNIDS	SKSNLYGQTSV	
gi-50913506.	NGTATYNARGASVYI	ALLTGATDGR	TYNPILLAASYN	GEGLVTKNIDS	SKSNLYGQTSV	
	130	140	150	160	170	180
	190	200	210	220	230	240
gi-50913506.	AKSSLP	SITKKVTGT	IIDDVNKKTT	SLGSVL	SYSLTFELPSYTKEAVNKTVVVSDNMSEG	
gi-50913506.	AKSSLP	SITKKVTGT	IIDDVNKKTT	SLGSVL	SYSLTFELPSYTKEAVNKTVVVSDNMSEG	
	190	200	210	220	230	240
	250	260	270	280	290	300
gi-50913506.	TFNFNSL	TVEWKGMAN	ITEDG	SVMVENT	KIGIAKEVN	NGFNL
gi-50913506.	TFNFNSL	TVEWKGMAN	ITEDG	SVMVENT	KIGIAKEVN	NGFNL
	250	260	270	280	290	300
	310	320	330	340	350	360
gi-50913506.	AVVNNKAIV	GEEGNPNKAEFFY	SNNPTKG	NTYDNLD	KKPDKGNGITS	SKEDSKIVYTYQIA
gi-50913506.	AVVNNKAIV	GEEGNPNKAEFFY	SNNPTKG	NTYDNLD	KKPDKGNGITS	SKEDSKIVYTYQIA
	310	320	330	340	350	360
	370	380	390	400	410	420
gi-50913506.	FRKVDSVSK	TPLIGAIFGVYDT	TSNK	LIDIVFT	NKNGYAISTQVSSGKY	KIKELKAPKGYS
gi-50913506.	FRKVDSVSK	TPLIGAIFGVYDT	TSNK	LIDIVFT	NKNGYAISTQVSSGKY	KIKELKAPKGYS
	370	380	390	400	410	420
	430	440	450	460	470	480
gi-50913506.	LNTETY	EITANWVTATV	KTSANSK	TTTYS	SDKNKATDN	SEQVGLKNGEFFYSIDSRTGN
gi-50913506.	LNTETY	EITANWVTATV	KTSANSK	TTTYS	SDKNKATDN	SEQVGLKNGEFFYSIDSRTGN
	430	440	450	460	470	480
	490	500	510	520	530	540
gi-50913506.	DVKEAY	TESTKALTD	GTTFSKSNEGSGT	VILLE	DIPNTKL	GELPSTGSIGTYLFKAIGSA
gi-50913506.	DVKEAY	TESTKALTD	GTTFSKSNEGSGT	VILLE	DIPNTKL	GELPSTGSIGTYLFKAIGSA
	490	500	510	520	530	540
	550					
gi-50913506.	AMIGAIGIYIV	KRRKA				
gi-50913506.	AMIGAIGIYIV	KRRKA				
	550					

FIGURE 58B

189/487

PCT/US2005/027239
gi-50913506.pep

/home/morama/gas/pili/align/orf84.pep

TRANSLATE of: orf84.seq check: 7868 from: 1 to: 2088
generated symbols 1 to: 696.

GETSEQ from morama, September 13, 2004 17:07.

SCORES Init1: 57 Initn: 83 Opt: 135 z-score: 75.1 E(): 1.2
 >>/home/morama/gas/pili/align/orf84.pep (696 aa)
 initn: 83 init1: 57 opt: 135 z-score: 75.1 expect(): 1.2
 Smith-Waterman score: 146; 24.4% identity in 262 aa overlap
 (232-462:316-567)

gi-50913506. 210 220 230 240 250
 KTTSLGSVLSYSLTFELPSYTKEAVNKTVVYSDNMSEGLTFNFNSLTVEWKGMAN---

orf84.pep EKEAVVDDAELKQINSEGQQEISVTWTNQLVTDE--KGMAVIYSVKEVDKNGELLEPKDY
 290 300 310 320 330 340

gi-50913506. 260 270 280 290 300 310
 ITEDGSVMVENTKIGIAKEVNNGFNLSFIY---DSLESISPNIKYAVVNNKAIVGEE
 orf84.pep IKKEDGLTVTNTYV---KPTSGHYDIEVTFGNGHIDITEDTPDI-VSGENQMKQIEGED
 350 360 370 380 390

gi-50913506. 320 330 340 350
 GNP-----NKAEFFYSNNP-----TKGNTYDNL-DKKP-DKGNGITTSKEDSKTVYTYQ
 orf84.pep SKPIDEVTEENLIEFGKNTMPGEEDGTSNKYEEVEDSRPVDTLSGLSSEQGQSGDMTIE
 400 410 420 430 440 450

gi-50913506. 360 370 380 390 400 410
 -----IAFRKVDSVSKTPLIGAIFGVYDTSNKLIDIVTTNKNGYAIISTQVSSGKYKIKE
 orf84.pep EDSATHIKFSKRD-IDGKELAGATMELRDSGGKTIS--TWISDGQVKDFYLMPGKYTFVE
 460 470 480 490 500 510

gi-50913506. 420 430 440 450 460 470
 LKAPKGYSLNTE-TYEITAN-WVTATVKTTSANSKSTTYTSDKNKATDNSEQVGWLKNGIF
 orf84.pep TAAPDGYESIATAITFTVNEQQVTVNGKATKGDAHTVMV-DAYKPTKGSGQVEDIEEKLP
 520 530 540 550 560 570

gi-50913506. 480 490 500 510 520 530
 YSIDSRPTGNDVKEAYIESTKALTDGTTESKSNEGSGTVILLETDIPNTKLGELPSTGSTG
 orf84.pep DEQGHSGSTTEIEDSKSSDVLIIGGQGQIVETTEDQTGMHGDSGCKTEVEDTKLVQSFHF
 580 590 600 610 620 630

gi-50913506.pep

/home/morama/gas/pili/align/gi-19745307.pep

gi|19745307|ref|NP_606443.1| protein F2-like protein [Streptococcus pyogenes MGA S8232]

SCORES Init1: 57 Initn: 83 Opt: 135 z-score: 75.1 E(): 1.2
 >>/home/morama/gas/pili/align/gi-19745307.pep (696 aa)
 initn: 83 init1: 57 opt: 135 z-score: 75.1 expect(): 1.2
 Smith-Waterman score: 146; 24.4% identity in 262 aa overlap
 (232-462:316-567)

	210	220	230	240	250	
gi-50913506.	KTTSLGSVLSYSLT	FELPSYTKEAVNKTVY	VSDNMSEGL	TNFNSLT	VEWKGMAN	
gi-19745307.	EKEVAVDDAELKQINSEGQQEISVTWTNQLVTDE	--KG	MAYIYSVKEVDKN	GELLEPKDY		
	290	300	310	320	330	340
gi-50913506.	ITEDGSVMVENT	KIGIAKEVNNGFNLSFIY	DSLESISP	NISYKAVVNNKAIVGEE		
gi-19745307.	IKKEDGLTVNTNTYV	--KPTSGHYDIEVTFGNGHIDIT	TEDTTPDI	-VSGENQM	KQIEGED	
	350	360	370	380	390	
	320	330	340	350		
gi-50913506.	GNP	NKAEFFYSNNP	TKGNTYDNL	-DKKP	-DKGNGITSKEDSKIVYTYQ	
gi-19745307.	SKPIDEVTENN	LIEFGKNTMPGEEDGTNS	MKYEEVEDSRP	VDTLSGLS	SEQGQSGDMTIE	
	400	410	420	430	440	450
	360	370	380	390	400	410
gi-50913506.	IAFRKVDSVSKTPLIGA	IFGVYDT	TSNKLIDIV	TTNKNGYAIST	QVSSGKYKIKE	
gi-19745307.	EDSAT	HIKFSKR	IDGKELAGATMELRDSSGKTIS	-TWISDGQVKDFY	LMPGKYTFVE	
	460	470	480	490	500	510
	420	430	440	450	460	470
gi-50913506.	LKAPKGYSLNTE	-TYBITAN	-WVTA	TVKTSANSKST	TYTSD	DKNKATDN
gi-19745307.	TAAPDGYEIATAIT	FTVNEQQ	QVTVNGKAT	KGDAHTIVMV	-DAYKPT	KGSQVIDIEEKLP
	520	530	540	550	560	570
	480	490	500	510	520	530
gi-50913506.	YSIDSRPTGNDV	KEAYTEST	KALTDGTTFSKSNEGSGT	VVL	ETDIP	PNTKLGELPSTGSIG
gi-19745307.	DEQGHSGSTTEIED	SKSSDVII	JGGQQQIV	ETTEDT	QTGMHGDS	GCKTEVEDT
	580	590	600	610	620	630

gi-50913506.pep
/home/moraima/gas/pili/align/gi-21909640.pep

gi|21909640|ref|NP_663908.1| protein F2 like fibronectin-binding protein [Streptococcus pyogenes MGAS315].

```

SCORES Init1: 56 Initn: 81 Opt: 134 z-score: 75.0 E(): 1.2
>>/home/morama/gas/pili/align/gi-21909640.pep (656 aa)
initn: 81 init1: 56 opt: 134 z-score: 75.0 expect(): 1.2
Smith-Waterman score: 156; 23.9% identity in 347 aa overlap
(232-547:202-524)

```

210	220	230	240	250	
gi-50913506.	KTTSLGSVLSYSLTFELPSYTKEAVNKTVYVSDNMSEGLTFNFNSLTVEWKGMAN				
	: : : : :				
gi-21909640.	EKEVAVDDAELKQINSEGQQEISVTWTNQLVTDE--KGMAITYSVIKEVDKNGELLEPKDY				
	180 190 200 210 220				
260	270	280	290	300	310
gi-50913506.	ITEDGSVMVENTKIGIAKEVNNGFNLNSFIY-----DSLESISPNIKYKAVVNNKAIVGEE				
	: : : : : : : : : : :				
gi-21909640.	IKKEDGLTVNTYV--KPTSGHYDIEVTFGNGHIDITEDTPDI-VSGENQMKQIEGED				
	230 240 250 260 270 280				

FIGURE 58D

PCT/US05/27239

gi-50913506. GNP-----NKAEFFYSNNP-----TKGNTYDNL-DKDP-DKGNGITSKEDSKIVYTYQ
 ::| | || ::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |:
 gi-21909640. SKPIDEVTEENNLLIEFGKNTMPGEEDGTNSNKYEEVEDSRPVDTLSGLSEQQQSGDMTIE
 290 300 310 320 330 340

gi-50913506. 360 370 380 390 400 410
 -----IAFRKVDSVSKTPLIGAIFGVYDTSNKLIDIVTTNKNGYAISTQVSSGKYKIKE
 ||| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |:
 gi-21909640. EDSATHIKFSKRD-IDGKELAGATMELRDSSGKTIS--TWLSDGQVKDFYLMPGKYTFVE
 350 360 370 380 390 400

gi-50913506. 420 430 440 450 460 470
 LKAPKGYSLNTE-TYEITAN-WVTATVKTANSKSTTYTSKDNKATDNSEQVGWLKNGIF
 ||| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |:
 gi-21909640. TAAPDGYESATAITFTVNEQQQVTVNNGATKGDAHIVMV-DAYKPTKGSGQV--
 410 420 430 440 450

gi-50913506. 480 490 500 510 520 530
 YSIDSRPTGNDVKEAYIESTKALTDTGTFSKSNEGSGTVLLETDIPNTKLGELPSTGSIG
 ||| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |:
 gi-21909640. -IDIEEKLPD-EQGHSGSTTEIEDSKSSDVIIIGGQGEVVDTTE--DTQSGMTGHSGST-
 460 470 480 490 500

gi-50913506. 540 550
 TYLFKAIGSAAMIGAIGIYIVKRRKA
 | :| :| :| :|:
 gi-21909640. TEIEDSKSSDVIIIGGQGEVVDTTEDTQSGMTGHSGSTTKIEDSKSSDVIVGGQQQIVETT
 510 520 530 540 550 560

gi-50913506.pep
 /home/morama/gas/pili/align/gi-28810263.pep

gi|28810263|dbj|BAC63201.1| protein F2-like protein [Streptococcus pyogenes SSI-1]

SCORES Init1: 56 Initn: 82 Opt: 134 z-score: 74.7 E(): 1.3
 >>/home/morama/gas/pili/align/gi-28810263.pep (733 aa)
 initn: 82 init1: 56 opt: 134 z-score: 74.7 expect(): 1.3
 Smith-Waterman score: 155; 23.6% identity in 347 aa overlap
 (232-547:316-638)

gi-50913506. 210 220 230 240 250
 KITSLGSVLSYSLTELEPSYTKEAVNKTIVVSDNMSEGLITFNENSLIVEWKGKMAN
 ||| |::| |::| |::| |::| |:
 gi-28810263. EKEVAVDPAELKQINSECQQEITSVITWNQEVLTDE-KGMAYIVSVKEVDKNCELLEPKDY
 290 300 310 320 330 340

gi-50913506. 260 270 280 290 300 310
 ITEDGSVMVENTKIGIAKEVNNGFNLIFIY--DSLESISPNTSYKAVVNNKAIVGEE
 | ::| ::| |::| |::| |::| |::| |::| |::| |::| |::| |:
 gi-28810263. IKKEDGLTVNTNYV--KPTSGHYDIEVTFGNGHIDITEDTPDVI-VSGENQMKQIEGED-
 350 360 370 380 390

gi-50913506. 320 330 340 350
 GNP-----NKAEFFYSNNP-----TKGNTYDNL-DKDP-DKGNGITSKEDSKIVYTYQ
 ::| | || ::| |::| |::| |::| |::| |::| |::| |::| |::| |:
 gi-28810263. SKPIDEVTEENNLLIEFGKNTMPGEEDGTNSNKYEEVEDSRPVDTLSGLSEQQQSGDMTIE
 400 410 420 430 440 450

360 370 380 390 400 410

FIGURE 58E

PCT/2005/27239

gi-50913506. IAFRKVDSVSKTPLIGAIFGVYDTSNKLIDIVTTNKNGYAISTQVSSGKYKIKE

|| | | : | | : | : | : | | : | : | : | : |

gi-28810263. EDSATHIKFSKRD-IDGKELAGATMELRDSSGKTIS--TWISDGQVKDFYLMPGKYTFVE

460 470 480 490 500 510

420 430 440 450 460 470

gi-50913506. LKAPKGYSLNTE-TYEITAN-WTATVKTTSANSKSTTYTSDKNKATDNSEQVGWLKNGIF

|| | | : | : | : | : | : | : | : | : | : |

gi-28810263. TAAPDGYEVATAITFTVNEQQVTVNGKATKGDAHIVMV-DAYKPTKGSGQV-----

520 530 540 550 560

480 490 500 510 520 530

gi-50913506. YSIDSRPTGNDVKEAYIESTKALTDGTTFSKSNEGSGTVLLETDPNTKLGELPSTGSIG

|| : | : | : | : | : | : | : | : | : | : |

gi-28810263. --IDIEEKLKD-EQGHSGSTTEIEDSKSSDVIICGGQGEVVDTTE--DTQSGMTGHSGST-

570 580 590 600 610 620

540 550

gi-50913506. TYLFKAIGSAAMIGAIGIYIVKRRKA

| : | : | : |

gi-28810263. TKIEDSKSSDVIVGGQQIVETTEDTQTMHGDSGRKTEVEDTKLVQSFHFIDNKEPESN

630 640 650 660 670 680

gi-50913506.pep

/home/morama/gas/pili/align/orf80.pep

TRANSLATE of: orf80.seq check: 9824 from: 1 to: 1056

generated symbols 1 to: 352.

GETSEQ from morama, September 13, 2004 17:11.

SCORES Init1: 45 Initn: 69 Opt: 113 z-score: 70.8 expect(): 2.1

>>/home/morama/gas/pili/align/orf80.pep (352 aa)

initn: 69 init1: 45 opt: 113 z-score: 70.8 expect(): 2.1

Smith-Waterman score: 123; 22.8% identity in 311 aa overlap

(284-556:49-352)

260 270 280 290 300 310

gi-50913506. KMANITEDGSVMVENTKIGIAKEVNNGFNLSFIYDSLESISPNIKYAVVN-NKAIVGEE

:: | : | : | : | : | : | : | : |

orf80.pep ATALGTASLNQNVAETAGVVTGKSLQVTKMTYDDEVLMPETAFTFTIEDDMTASCKE

20 30 40 50 60 70

320 330 340 350 360 370

gi-50913506. GNPN-KAEFFYSMNPTKGNTYDNLDKKPDKGNGTTSKEDSKIVVYTYQIAFRKVDSVSKTP

| : | : | : | : | : | : | : | : | : | : |

orf80.pep GSLDIKNGLVEGLDKQVTVKYKNTDKTSQKTK-TAQFDFSKVKEPAIGVYRYMVSEKNDK

80 90 100 110 120 130

380 390 400 410 420

gi-50913506. LIGAIFGVYDTSNKLIDIVTTNKNG-----YAISTQ-VSSGKYKIKELKAPKGYSLN

| : | : | : | : | : | : | : | : | : | : |

orf80.pep KDGI---TYDDKKWTVDVYVGNKANNEEGFEVLYIVSKEGTSSTKPIEFTNSIKTTSIK

140 150 160 170 180 190

430 440 450 460

gi-50913506. TETYEITANW-----VTATVKTTSANSKSTTYTSDKNKATDNSEQVG----WLKNGL

| : | : | : | : | : | : | : | : | : |

orf80.pep IEK-QITGNAGDRKKSFNFTLTLQPSEYYKTGSVVKIEQDGSKKDTIGTPYKFTLGHGK

200 210 220 230 240 250

470 480 490 500 510

FIGURE 58F

PCT/US05/27239

gi-50913506. FYSIDSRPTGNDV-----KEAYI-----ESTKALTDGTTFSKSNEGSGTVLLETDI
 orf80.pep SVMLSKLPIGINYYLSEDEANKDGYTTATLKEQGKEKSSDFTLSTQNQKTDESADEIVV
 260 270 280 290 300 310

gi-50913506. PNTKLGELPSTGSIGTYLFKAIGSAAMIGAIGIYIVKRRKA
 orf80.pep TNKRDTQVP-TGVVGT LAPFAVL SIV AIGGV-IYITKRKKA
 320 330 340 350

gi-50913506.pep

/home/morama/gas/pili/align/gi-19224137.pep

gi|19224137|gb|AAL86408.1|AF447492_5 EftLSL.A [streptococcus pyogenes]

SCORES Init1: 45 Initn: 69 Opt: 109 z-score: 69.8 E(): 2.4
 >>/home/morama/gas/pili/align/gi-19224137.pep (342 aa)
 initn: 69 init1: 45 opt: 109 z-score: 69.8 expect(): 2.4
 Smith-Waterman score: 169; 26.0% identity in 334 aa overlap
 (257-556:25-342)

230 240 250 260 270 280
 gi-50913506. NKTVYVSDNMSEGGLTFNFNSLTVEWKGMANITEDGSVMVENTKIGIAKEVNNGFNL SFI
 | : : : | : : | : : | : : | : | : |
 gi-19224137. MKKNKLLLATAILATALGTASLNQNVKAETAGVVSSGQLTIKK SITN-FN---
 10 20 30 40

290 300 310 320 330 340
 gi-50913506. YDSLESISPNSIYKAVVN-NKATVGEEN-PNKAEEFYSSNNPTKGNTYDNLDKPKDGNG
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 gi-19224137. DDTL--LMPKTDYTFSVNPDSAATGTESNLPIKPGIAVNQDIK-VSYSNTDKTSGKEQ
 50 60 70 80 90 100

350 360 370 380 390 400
 gi-50913506. ITSKEDSKIVYTQIAFRKVDSVSKTPLIGAIFGV-YDTSNKLIDIVTTNKNGYAISTQV
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 gi-19224137. VVV-DFMKVTFPSVGIYRYVVTENK---GTAEGVTDYDDTKWLVDVYVGNNNEKGGLPEKY
 110 120 130 140 150 160

410 420 430 440 450
 gi-50913506. SSGKYKTIKELKAPKGY-SLNTETYEITANWVATVKT SANSKSTTYTS DKNKAT DNS--
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 gi-19224137. IVSKKGDSATKEPIOFNNSETTS LKIEKE-VTGNTGDHKKAFTTLQPN EYYEASSV
 170 180 190 200 210 220

460 470 480 490
 gi-50913506. --EQVWLKN--GIFYST--DSR----PTGND--VKEAYIE---STKALT DG
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 gi-19224137. VKIEENGQTKDVKIGEAYKFTLNDQS VILSKLPVGINYKVEEAANQGGYTTATLKDG
 230 240 250 260 270 280

500 510 520 530 540 550
 gi-50913506. TTFSKSNEG---SGTVLLETDI PNTKLGELPSTGSIGTYLFKAIGSAAMIGAIGIYIVK
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 gi-19224137. EKLSTYNLGQEHKTDADEIVVTNNRDTQVP-TGVVGT LAPFAVL SIV AIGGV-IYITK
 290 300 310 320 330

gi-50913506. RRKA

|:|

FIGURE 58G

PCT/US05/27239
gi-19224137. RKKA
340

194/487

gi-50913506.pep
/home/morama/gas/pili/align/gi-19224141.pep

gi|19224141|gb|AAL86412.1|AF447492_9 protein F2 [Streptococcus pyogenes]

```
SCORES Init1: 73 Initn: 73 Opt: 118 z-score: 68.9 E(): 2.7  
>>/home/morama/gas/pili/align/gi-19224141.pep (1161 aa)  
initn: 73 init1: 73 opt: 118 Z-score: 68.9 expect(): 2.7  
Smith-Waterman score: 174; 23.9% identity in 406 aa overlap  
(115-483:277-645)
```

90	100	110	120	130	140
gi-50913506.	VSLTETGPTSQEITTIANGINTGKIKPFSTENVSISNGTATYNARGASVY--IALLTGAT				
		: : :::: : : :			
250	260	270	280	290	300
gi-19224141.	IYTFTDYIAGLDKVQLSAELSLFLENKEVLENTSISMFKSTIGGQEITYKGTVNVLYGNE				

	150	160	170	180	190	
gi-50913506.	DGRTYNPILLAASYNNEG	NLVTKNIDS	KSNLYGQTSVAKSLPSITKKVTG			T
gi-19224141.	STKESNYITNGLSNVG-GSIESYNTETGEFVWYVYVNPNRTNIPYATMNLWGFRARSNT					
	310	320	330	340	350	360

260 270 280 290 300 310
 gi-50913506. WKGKMANITEDGSVMVENTKIGIAKEVNNGFNLSFIYDSLESI SPNIS YKAVVNNKAIVG
 ||:: :| ::|| | :|| : :| | :| .:| ::||:
 gi-19224141. -----MTKRQRIDFG--NNIQNKAFI IKV--TGKTDQSGKPLVVQSNLAS
 420 430 440 450 460

320	330	340	350	360
gi 50913506 : EEGNPNAEFFYSNNPTKGNTY--DNLDKKPDKGNGITSKEDSKIVYTY				QIAF
gi 19224141 : FRGASEYAAF---TPVGGNVYFQNEIALSPSKGSGSKSEFTKPSITVANLKRVAQLRF				
470	480	490	500	510

370 380 390 400 410 420
gi-50913506. RKVDSVSKTPLIGAIFGVYDTSNKLIDI-VTTNKNGYAISTQVSSGKYKIKELKAPKGYS
: : | : ::::| | | : ::::: : : : | : | : ::::| | | : | |||||:
gi-19224141. KKM-STDNVPLPEAAFELRSSNGNSQKLEASSNTQGEVFKDLSGTYDLYETKAPKGYQ
520 530 540 550 560 570

gi-50913506.	470	480	490	500	510	520
	GIFYSIDS RPTGNDVKEAYIESTKALT DGTFSKSNEGSGTVLLET DIPN TKL GELPSTG					
	: : : :					
gi-19224141.	----RPDQRPAKIQVQLLQNGQKMPNQIQEVTKDNDWSYHFKDLPKYDAKNQEYKYSVEE					
	640	650	660	670	680	

FIGURE 58H

gi-50913506.pep

/home/morama/gas/pili/align/gi-21909636.pep

gi|21909636|ref|NP_663904.1| conserved hypothetical protein [Streptococcus pyogenes MGAS315]

SCORES Init1: 45 Initn: 98 Opt: 96 z-score: 66.1 E(): 3.8
 >>/home/morama/gas/pili/align/gi-21909636.pep (344 aa)
 initn: 98 init1: 45 opt: 96 Z-score: 66.1 expect(): 3.8
 Smith-Waterman score: 181; 25.3% identity in 312 aa overlap
 (298-556:44-344)

270	280	290	300	310	320
gi-50913506.	NTKIGIAKEVNNGFNL	SFIYDSLESISPN	SYKAVVNNKAIVGEEGNPNKAEFFYSNNPT		
	20	30	40	50	60
gi-21909636.	ATALGTASLNQNVKAETAGVSENAKLIVKKT	FDSYTDNEVLMPKADYTFKVE	--ADSTA		
	80	90	100	110	120

330	340	350	360		
gi-50913506.	KGNTYDNL	DKKPDKGNGIT	SKEDSKIVYTYQIAFRKV		DSVSK
	130	140	150	160	170
gi-21909636.	SGTKDGLEIKPGIVNGLTEQIISYTNTDKPDSKV	KST-EFDFSKVVFPGIGVYRYTVSE			
	80	90	100	110	120

370	380	390	400	410	420
gi-50913506.	TPLIGAIFGV-YDTSNKLIDIVTTNKNGY	AI	STQVSSGKYKIKELKAPGY	--SLNTETY	
	130	140	150	160	170
gi-21909636.	KQ--GDVEGITYDTKKWTV	DVYVG	NKEGGFEPKFIVSKEQGTDVKKPVNFNNNSFATT	SL	
	80	90	100	110	120

430	440	450	460		
gi-50913506.	EITANWVTATVKTSAN	SKSTTYTSDK	NKATDNSE	QVGW	--LKNGI
	190	200	210	220	230
gi-21909636.	KVKKNVSGNTGELQKEFDFTLTLNESTNFKKDQIVS	LQKGNEKFEVKIGTPYKFKLKNGE			
	80	90	100	110	120

470	480	490	500	510	
gi-50913506.	FYSIDSRTGNDVKEAYIESTKALT	DGTTFSKS	NEGSGTVLL	--ETD	
	190	200	210	220	230
gi-21909636.	SIQLDKLPVGITYKV	NEMEANK	DGYKTTASLKEGDGQSKMYQLDMEQKTDESAD	EIV	
	80	90	100	110	120

520	530	540	550		
gi-50913506.	IPNTKLGELPSTGSIGTYLFKAIGSAAMIGAIGIYIVKRRKA				
	190	200	210	220	230
gi-21909636.	VTNKRDQTQVP-TGVVGT	LAPFAVLSIVAIGGV	IYITKRKKA		
	80	90	100	110	120

gi-50913506.pep

/home/morama/gas/pili/align/gi-28810259.pep

gi|28810259|dbj|BAC63197.1| hypothetical protein [Streptococcus pyogenes SSI-1]

SCORES Init1: 45 Initn: 98 Opt: 96 z-score: 66.0 E(): 3.8
 >>/home/morama/gas/pili/align/gi-28810259.pep (349 aa)
 initn: 98 init1: 45 opt: 96 Z-score: 66.0 expect(): 3.8
 Smith-Waterman score: 181; 25.3% identity in 312 aa overlap
 (298-556:49-349)

PCT/US05/227239

	270	280	290	300	310	320
gi-50913506.	NTKIGIAKEVNNGFNLFSIYDSLES-ISPNIKYAVVNNKAIVGEEGNPKAEFFYSNNPT					
gi-28810259.	ATALGTASLNQNVKAETAGVSENAKLIVKKTFSYTDNEVLMPKADYTFKVE--ADSTA	20	30	40	50	60
						70
	330	340		350	360	
gi-50913506.	KGNTYDNLDKKPDKGNGIT-----SKEDSKIVYTYQIAFRKV-----DSVSK	80	90	100	110	120
gi-28810259.	SGKTKDGLEIKPGIVNGLTEQIISYTNTDKPSDKVKST-EFDFSKVVFPGLGVYRYTVSE					130
	370	380	390	400	410	420
gi-50913506.	TPLIGAIFGV-YDTSNKLIDIVTTNKNGYAISTQVSSGKYKIKELKAPKGY--SLNTETY	140	150	160	170	180
gi-28810259.	KQ--GDVEGITYDTKKWTVDVYVGNGKEGGFEPKFIVSKEQGTDVKKPVNFNNSFATTSL					190
	430	440		450	460	
gi-50913506.	EITANWVTATVKTSAN-----SKSTTYTSDK---NKATDNSE-QVGW-----LKNGI	200	210	220	230	240
gi-28810259.	KVKKNVSGNTGELOKEFDFTLTNESTNFKDQIVSLQKGNEKFPEVKIGTPYKFKLKNGE					250
	470	480	490	500	510	
gi-50913506.	FYSIDSRTGNDVKEAYIESTKALTDTGTTESKS-NEGSGTVLL-----ETD-----	260	270	280	290	300
gi-28810259.	SIQLDKLPVGITYKVNEMEANK--DGYKTTASLKEGDGQSKMYQLDMEQKTDEADEIV					
	520	530	540	550		
gi-50913506.	IPNTKLGELPSTGSIGTYLFKAIGSAAMIGAIGIYIVKRRKA	310	320	330	340	
gi-28810259.	VTNKRDTQVP-TGVVGTAPFAVLSIVAIGGV-IYITKRKKA					
gi-50913506.pep						
/home/morama/gas/pili/align/gas15.pep						
GAS15 GAS15						
SCORES Init1: 42 Initn: 68 Opt: 96 z-score: 63.8 E(): 5						
>>/home/morama/gas/pili/align/gas15.pep (762 aa)						
initn: 68 init1: 42 opt: 96 z-score: 63.8 expect(): 5						
Smith-Waterman score: 96; 23.4% identity in 269 aa overlap						
(283-535:222-470)						
	260	270	280	290	300	
gi-50913506.	GKMANITEDGSVMVENTKIGIAKEVNNGFNLFSIYDSLES-ISPNIKYAVVNNKAIVGEEGNPKAEFFYSNNPTKGN	200	210	220	230	240
gas15.pep	VWYYSDNAPISNPDESFKRESESNLVSTSQSLMRQALKQLIDPNLATKMPKQVPDDFOL					250
	310	320	330	340	350	360
gi-50913506.	AIVGEEGNPKAEFFYSNNPTKGN	260	270	280	290	300
gas15.pep	YDNLDKKPDKGNGITSKEDSKIVYTYQIAFRKV-V					
	SIFESEDKDKYNKGYQNLLSGGLVPT-KPPTPGDPPMPPNQPO--TTSVLIRKYAT	370	380	390	400	410
						420

FIGURE 58J

PCT Zusatz 222 235

```

gi-50913506. DSVSKTPLIGAIFGVY-DTSNKL-IDIVTTNKNGYAISTQVSSGKYKIKELKAPKGYSLN
          : ||| | ||| : : | : ||| : : ||| | | | : ||| : ||| : ||| : ||| : ||| : |
gas15.pep   GDYSKL-LEGATLQLTGDNVNSFQARVFSSNDIGERI--ELSDGTYTLTELNSPAGYSIA
          310      320      330      340      350      360

```

```

gi-50913506. TE-TYEITANWVTATV--KTSANSKSTTYTSDKNKATDNSEQVGWLKN---GIFYSIDSR
| : :: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
gas15.pep EPITFKVEAGKVYTIDGKQIENPNKEIVEPYVSVEAYNDFEEFSVLTQNYAKFYYYAKNK
370      380      390      400      410      420

```

gi-50913506.	480	490	500	510	520	530
	PTGNDVKEAYIESTKALT	DGTTFSKSNEGSGTVLLET	DIPNTKL	GELPSTGSIGTYLFKA		
gas15.pep	:::: : : :	:: : :	: :			
	NGSSQVVYCFNADLKSPD	-----	SEDDGGKTMT	-----	PDFTTGEVKYTHIAGRDLFKY	
	430	440	450	460	470	

540 550

qi-50913506. IGSAAAGATGATGTTVKRRKA

gas15.pep TVKPRDTDPTFLKHIIKKVIEKGYREKGQAIYEYSGLTETQLRAATQLATYYFTDSAELDK
480 490 500 510 520 530

gi-50913506.pep

/home/morama/gas/pili/align/gi-13621428.pep

gi|13621428|gb|AAK33238.1| hypothetical protein [Streptococcus pyogenes]

```

SCORES Init1: 41 Initn: 41 Opt: 87 z-score: 63.6 E(): 5.2
>>/home/morama/gas/pili/align/gi-13621428.pep (340 aa)
initn: 41 init1: 41 opt: 87 Z-score: 63.6 expect(): 5.2
Smith-Waterman score: 109; 22.6% identity in 345 aa overlap
(256-556:17-340)

```

230 240 250 260 270 280
gi-50913506. VNKTIVVSDNMSEGLTNFNSLTVEWKGMANITEDGSVMVENTKIGIAKEVNNGFNLSF
| | | : : | : : | : ||:
gi-13621428. MKLRHLLLTTGAALTSEFAATTVHGTVVNGAKLTVTK-----NLDD

290. 300. 310. 320. 330. 340.
gi_50913506. IYDSELSISPNTSYKAVVNNKAIVGEEGNPNAEFFYSNNPTKGNTYDNLDKKPDKGNGL
gi_13621428. VNSN- ALIPPTDFTKIEEDITVNEEGNKFRGVAL- NTPMTKVTYTNSDK- GGSNTIK

350 360 370
gi-50913506. TSKED-SKI-----VYTYQTAFRKVDSV-----SKTPLIGAIF
| : | | : | | : | | : | | : | | : | | : | | : | |
gi-13621428. TAEFDSEVTFEKPGVYYYYKVTEEKIDKVPGVSYDTSYTVQHVWLNEEQKPVATYIV

gi-50913506. GYVDTSNKLIDIVTT-NKNGYAISTQVS-SGKYKIKELKAPKGYSINTETYEITANWTA
| : | : | :: :: : :: : || : | : | :: : | : | :: : | : | :: : |
gi-13621428. GYKEGSKVPIQFKNSLDSTTLTVKKVSGTGGDRSKDFNE--GLTLKANOYYKASEKVM

440 450 460 470 480 490
gi-50913506. TVKTSANSKSTTYTSDKNKATDNSEQVGWLKNGIFYSIDSRPTGND--VKEAYIESTKAL

FIGURE 58K

PCT/US05/23230
gi-13621428. E-KTTKGQAPVQT---EASIDQLYHFTLKDGESIKVTNLPGVVDYVVTEDDYKSEKYT
220 230 240 250 260

500 510 520 530 540
gi-50913506. T-----DGT-----FSKSNEGSGTVILLETDIPNTKLGELPSTGSIGTYLFKAIGSAA
| |::| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
gi-13621428. TNVEVSPQDGAVKNIAGNSTEQETSTDKDMTITFTNKKDFEVPTGVAMTVAPYIALGIVA
270 280 290 300 310 320

550
gi-50913506. MIGAIGIYIVKRRKA
: | :| :| :| :|
gi-13621428. VGGAL--YFVKKKNA
330 340

gi-50913506.pep
/home/morama/gas/pili/align/gi-19224135.pep

gi|19224135|gb|AAL86406.1|AF447492_3 Cpa [Streptococcus pyogenes]

SCORES Init1: 41 Initn: 41 Opt: 86 z-score: 61.0 E(): 7
>>/home/morama/gas/pili/align/gi-19224135.pep (756 aa)
initn: 41 init1: 41 opt: 86 z-score: 61.0 expect(): 7
Smith-Waterman score: 101; 19.6% identity in 306 aa overlap
(243-535:193-462)

220 230 240 250 260 270
gi-50913506. SLTFELPSYTKEAVNKTVYVSDNMSEGLTFNFNSLTVEWKGKMANITEDGSVMVENTKIG
: : : :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
gi-19224135. PKNANGYMDKIEPLNAILVTQQAVWYYSDDSYGNIKTLWASEL---KDGKIDFEQVKL-
170 180 190 200 210
280 290 300 310 320 330
gi-50913506. IAKEVNNNGFNLSFIYDSLESISPNISYKAVVNNKAIVGEEGNPN--KAEFFYSNNPTKGN
:| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
gi-19224135. -MREAYS---KLISDDLEETSKNKLPQGSKLNIFVPQDKSVQNLLSAEYVPESPAPGQ
220 230 240 250 260 270
340 350 360 370 380
gi-50913506. TYDNLDKPKDKNGNITSKEDSKIVYTY-QIAFRKVDSVSKTPLIGAIFGVYDTSNKLIDI
:| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
gi-19224135. S---PEPP---VQTKKTSVITRKYAEFDYSKILLEGATLRLTGE-DILDFQEKG---V
280 290 300 310
390 400 410 420 430 440
gi-50913506. VTTTNKNGYAIISTQVSSGKYKIKELKAPKGYSINTET-YEITANWTATVKTTSANSKSTT-
:| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
gi-19224135. FQSNGTGEGKI--ELSNGTYTLTETSSPDGYKIAEPIKFRVVNKKFIVQKDGSOVENPNK
320 330 340 350 360 370
450 460 470 480 490
gi-50913506. ---YTSOKNATDNSEQVG---WLKNGIFYSIDSRPTGNDVKEAYTESTKALTIDGTTF
:| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
gi-19224135. EVAEPYSVEAYSDMQDSNYINPETFTPYGFYYAKNKKDKSSQVYCFN-----ADLHSP
380 390 400 410 420 430
500 510 520 530 540 550
gi-50913506. SKSNEGSGTVILLETDIPNTKLGELPSTGSIGTYLFKAIGSAA MIGAIGIYIVKRRKA
:| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
gi-19224135. PESEDGGGTI--DPDISTMK--EVKYTHTAGSDLFKYALRPRDTNPEDFLKHIKKVIEKG
440 450 460 470 480

FIGURE 58L

gi-19224135. YNKKGDSYNGLTETQFRAATQLATYYFTDSTDLKTLKTYNNGKGYHGFESMDEKTLAVTK
 490 500 510 520 530 540

gi-50913506.pep
 /home/morama/gas/pili/align/gi-50913505.pep

gi|50913505|ref|YP_059477.1| Collagen adhesion protein [Streptococcus pyogenes M GAS10394]

SCORES Init1: 70 Initn: 149 Opt: 86 z-score: 60.1 E(): 7.8
 >/home/morama/gas/pili/align/gi-50913505.pep (1036 aa)
 initn: 149 init1: 70 opt: 86 Z-score: 60.1 expect(): 7.8
 Smith-Waterman score: 120; 21.5% identity in 469 aa overlap
 (33-428:503-966)

10 20 30 40 50 60
 gi-50913506. NRRETVREKILITAKKMLACLAIALAVVGLGMTRVS-ALSKDDTAQLKITNIEGGPTVTL
 |::||:::||:::||:::||:::||
 gi-50913505. YSQWVYHGTFDYSSYQGESFNRGQIHYRYRGIVSVSDGIRRDAVKNSLLGVNG----L
 480 490 500 510 520

70 80 90 100 110
 gi-50913506. YKIGEGVYNTNGDSFINFK---YAEGVSLTETGPTSQEIT-TIANGINTGKIKPFSTEN
 ::||::||:::||:::||:::||:::||:::||:::||:::||:::||:::||:::||:::||:::||:
 gi-50913505. LQRFVNINPENKLSVIGFQGSADYHAGKWYPDQSPRGGFYQPNLNNSRDAELLKGWSTNS
 530 540 550 560 570 580

120 130 140 150 160
 gi-50913506. VSISNGTATYNARGASVYIALLTGAT-----DGRTYNPILLAAS-YNGEGNLVTK
 ::||:::||:::||:::||:::||:::||:::||:::||:::||:::||:::||:::||:::||:
 gi-50913505. LLDPNTLTALHNNGTNYHAALLKAKEILNEVKDDGRRKIMIFISDGVPFYFGEDGYRSG
 590 600 610 620 630 640

170 180 190 200 210
 gi-50913506. NIDS--KSNYLYGQ--TSVA---KSSLPSITKKVTGTTDDVNKKTTSLGSVLSYSLTFE
 ||::||:::||:::||:::||:::||:::||:::||:::||:::||:::||:::||:::||:
 gi-50913505. NGSSNDRNNVTRSQEWSKLAIDEFKARYPNLSIYSLGVSKDINSDTASSSPVVLKYLSE
 650 660 670 680 690 700

220 230 240 250 260
 gi-50913506. IIPSY-----TKEAVNKIVY-----VSDNMSEGTF-NFNSLTVEWKGRMANITE
 ::||:::||:::||:::||:::||:::||:::||:::||:::||:::||:::||:::||:
 gi-50913505. EHYYGITDTAELEKTLNKIVEDSKSQLGISDSLSQLVYDYYDKQPDVILVTRSKVNDETE
 710 720 730 740 750 760

270 280 290
 gi-50913506. -----DGSVMVENTKIGIAKEVNNGFNLNSFIYDS-LESIISPNT-----
 ||::||:::||:::||:::||:::||:::||:::||:::||:::||:::||:
 gi-50913505. ILYQKDQVQEAGKDIIDKVVFPTKTTSQPKGVTLTFKSDYKVDEYTTLSFNVKASDE
 770 780 790 800 810 820

300 310 320 330 340 350
 gi-50913506. SYKAVVNNKAIVGEEGNPNKAEFFYSNNPTKGNTYDNLDDKKPDK-GNGITSKEDSK-----
 ::||:::||:::||:::||:::||:::||:::||:::||:::||:::||:
 gi-50913505. AYEKYKDNEGRYSEMGSDDTDYGTNQTSSGKGGLPSNSDASVNYMADGREQKLPYKHPVI
 830 840 850 860 870 880

360 370 380 390 400
 gi-50913506. IVYTYQTAFRKVDS-VSKTPLIGAIFGVYDTSNKLI-DIVTTNKNGYAIISTQVSSGK-YK
 ||::||:::||:::||:::||:::||:::||:::||:::||:::||:

FIGURE 58M

PCT/US05/27239

gi-50913505. QVKTVPITFTKVDADNNQKKLAGVEFELRKEDKKIVWEKTTGSNGQLNFKYLQKGKTYY
 890 900 910 920 930 940

gi-50913506. IKELKAPKGYSLNTEETYEITANWVTATVTSANSKSTTYTSDKNKATDNSEQVGWLNGI
 : | || ||:: : :::::
 gi-50913505. LYETKAALKGYTLPEPNPWEVAVANNGDIKVKHPIEGELKSKDGSYMTKNYKTYQLPSSGR
 950 960 970 980 990 1000

gi-50913506.pep

/home/morama/gas/pili/align/gi-13621430.pep

gi|13621430|gb|AAK33240.1| hypothetical protein [Streptococcus pyogenes]

SCORES Init1: 43 Initn: 67 Opt: 67 z-score: 59.2 E(): 8.7
 >>/home/morama/gas/pili/align/gi-13621430.pep (215 aa)
 initn: 67 init1: 43 opt: 67 z-score: 59.2 expect(): 8.7
 Smith-Waterman score: 67; 27.8% identity in 90 aa overlap
 (433-522:60-143)

410 420 430 440 450 460
 gi-50913506. VSSGKYKIKELKAPKGYSLNTEETYEITANWVTATVTSANSKSTTYTSDKNKATDNSEQV
 : | || ||:: : :::::
 gi-13621430. TASINIEVINQVDVATNKQSSDIDETFMFVIEALDKESPLPNSVT-TSVKGNGKTSFEQL
 30 40 50 60 70 80

470 480 490 500 510 520
 gi-50913506. GWLKNGIFYSIDSRTPTGNDVKEAYIESTKALTDGTTFSKSNEGSTVLLIEDIPNTKLGE
 : : | :: : | :: : | :: : : | : | : | : | : | : | : | : | : | : | : | : | : |
 gi-13621430. TFSEVGQYHYKIHQILLGKNSQYHYDETVEYEVVIYVLY---NEQSGA--LETNLVSNKLGE
 90 100 110 120 130 140

530 540 550
 gi-50913506. LPSTGSIGTYLFKAIGSAAMTGAIGIYIVKRRKA

gi-13621430. TEKSELIFKQEYSEKTPEPHQPDTEKEKPKQRNGILPSTGEMVSÝVSALGIVLVATIT
 150 160 170 180 190 200

gi-50913506.pep

/home/morama/gas/pili/align/gi-19745303.pep

gi|19745303|ref|NP_606439.1| hypothetical protein [Streptococcus pyogenes MGAS82
 321]

SCORES Init1: 51 Initn: 106 Opt: 69 z-score: 58.4 E(): 9.5
 >>/home/morama/gas/pili/align/gi-19745303.pep (344 aa)
 initn: 106 init1: 51 opt: 69 z-score: 58.4 expect(): 9.5
 Smith-Waterman score: 129; 24.0% identity in 308 aa overlap
 (298-556:44-344)

270 280 290 300 310 320
 gi-50913506. NTKIGIAKEVNNGFNLSFIYDSLESISPNISYKAVVNNKAIVGEEGNPNKAEFFYSNNPT
 : : : :: : | : : : | : | : : :
 gi-19745303. ATALGTASLNQNVAETAGVIDGSTLUVKKTFPSYTDDKVLMPKADYTFKVE---ADDNA
 20 30 40 50 60 70

330 340 350 360 370
 gi-50913506. KGN TYDNL DKK PDK GNGI-----TSKEDSKIVYTYQIAFRKVD SVSKT--P
 ||:| |:|| || :|:| :|:| | :| | :| |

201/487

gi-197433037 KGKTKDGLDKPGVLDGLENTKTIHGNSDKTTAKEKSVNFDANVKFPGVGVYRTVSE
80 90 100 110 120

80 90 100 110 120 130

380 390 400 410 420
gi-50913506. LIGAIFGV-YDTSNKLIDIVTTNKN--GYAISTQVSS-GKYKIKELKAPKGYSINTETYE

gi-19745303. VNGNKAGIAYDSQQWTVDVYVNVNREDGGFEAKYIVSTEGGQSDKKPVLFKNF-FDTTSLK
140 150 160 170 180

430 440 450 460 470
gi-50913506. ITANWVTATVKTSANSKTTYTSDKNKATDNSEOVGWLKNGIF-----YSIDSP-----

gi-19745303 VTKK-VTGNTGEHQRSFSFTLLLTPNECFEKQGVVNILQGGETKKVVIGEYSFTLKDKE
190 200 210 220 230

480 490 500 510

gi-197453-03. SVTLSQLPGIEYKVTEEDVTKGYKTSATLKDGFVTDGYNLGDKSTTDESMDELYTWW

520 530 540 550

gi-50913506. KLGEPLSTGSIGTYLFKAIGSAAMIGAIGIYIVKRRKA
1674738

310 320 330 340

! Distributed over 1 thread

! Distributed over 1 thread.
! Start time: Wed Sep 15 18:45:54 2004
! Completion time: Wed Sep 15 18:46:02 2004

CPU time used: Database scan: 0:00:00.1

```
Database scan: 0:00:00.1
Post-scan processing: 0:00:01.9
Total CPU time: 0:00:02.0
```

! Output File: gi-50913506.fasta

FIGURE 580

FIGURE 580

! !SEQUENCE_LIST 1.0505 / 27235

(Peptide) FASTA of: gi|13621430|gb|AAK33240.1| hypothetical protein [Streptococcus pyogenes]

TO: *.pep Sequences: 56 Symbols: 22,803 Word Size: 2

Scoring matrix: GenRunData:blosum50.cmp

Variable pamfactor used

Gap creation penalty: 12 Gap extension penalty: 2

Histogram Key:

Each histogram symbol represents 1 search set sequences

z-scores computed from opt scores

z-score	obs	exp
(=)	(*)	

< 20	0	0:
22	0	0:
24	0	0:
26	0	0:
28	0	0:
30	0	0:
32	0	0:
34	0	1:*
36	0	2: *
38	0	3: * *
40	0	4: * *
42	0	5: * *
44	1	5:= * *
46	7	5:====*==
48	21	5:====*=====
50	7	5:====*==
52	2	4:== * *
54	4	3:==*=
56	3	3:==*
58	0	2: *
60	4	2:==*==
62	1	2:==*
64	0	1:*
66	0	1:*
68	0	1:*
70	0	1:*
72	0	0:
74	0	0:
76	0	0:
78	0	0:
80	3	0:====
82	2	0:==
84	0	0:
86	0	0:
88	0	0:
90	0	0:
92	0	0:
94	0	0:
96	0	0:
98	0	0:
100	0	0:

FIGURE 59

102 P00T/LIG05/27239
 104 0 0:
 106 0 0:
 108 0 0:
 110 0 0:
 112 0 0:
 114 0 0:
 116 0 0:
 118 0 0:
 >120 1 0:=

Joining threshold: 36, opt. threshold: 24, opt. width: 16, reg.-scaled

The best scores are:

	init1	initn	opt	z-sc E(55)..
/home/morama/gas/pili/align/gi-13621430.pep	Begin: 1	End: 215		
! gi 13621430 gb AAK33240.1 hypothetical	1338	1338	233.9	1.8e-09
/home/morama/gas/pili/align/gi-19745305.pep	Begin: 1	End: 193		
! gi 19745305 ref NP_606441.1 hypothetical	163	243	273	82.2 0.5
/home/morama/gas/pili/align/gi-28810261.pep	Begin: 2	End: 187		
! gi 28810261 dbj BAC63199.1 hypothetical	164	239	268	81.5 0.55
/home/morama/gas/pili/align/gi-19224139.pep	Begin: 2	End: 187		
! gi 19224139 gb AAL86410.1 AF447492...	164	236	265	81.0 0.57
/home/morama/gas/pili/align/orf82.pep	Begin: 30	End: 222		
! TRANSLATE of: orf82:seq check: 4296...	163	235	264	81.0 0.58
/home/morama/gas/pili/align/gi-21909638.pep	Begin: 2	End: 181		
! gi 21909638 ref NP_663906.1 hypothetical	164	239	261	80.5 0.62
/home/morama/gas/pili/align/gi-19745303.pep	Begin: 84	End: 183		
! gi 19745303 ref NP_606439.1 hypothetical	121	121	126	61.4 6.7
/home/morama/gas/pili/align/gi-13621428.pep	Begin: 6	End: 174		
! gi 13621428 gb AAK33238.1 hypothetical	58	86	122	60.9 7.2
/home/morama/gas/pili/align/gi-19224137.pep	Begin: 93	End: 201		
! gi 19224137 gb AAL86408.1 AF447492...	88	88	119	60.4 7.5
/home/morama/gas/pili/align/gi-50913503.pep	Begin: 549	End: 625		
! gi 50913503 ref YP_059475.1 Fibron...	73	73	117	60.4 7.6
/home/morama/gas/pili/align/gi-19224134.pep	Begin: 631	End: 697		
! gi 19224134 gb AAL86405.1 AF447492...	73	73	115	60.1 7.8

\End of List

gi-13621430.pep

/home/morama/gas/pili/align/gi-13621430.pep

gi|13621430|gb|AAK33240.1| hypothetical protein [Streptococcus pyogenes]

SCORES Init1: 1338 Initn: 1338 Opt: 1338 z-score: 233.9 E(): 1.8e-09
 >>/home/morama/gas/pili/align/gi-13621430.pep (215 aa)
 initn: 1338 init1: 1338 opt: 1338 z-score: 233.9 expect(): 1.8e-09
 Smith-Waterman score: 1338; 100.0% identity in 215 aa overlap
 (1-215:1-215).

	10	20	30	40	50	60
gi-13621430.	MKK	SILRILAIGYLLMSFC	LDSVEAENLTASINIEV	I	QVDVATNKQSSDIDET	TMFVI
gi-13621430.	MKK	SILRILAIGYLLMSFC	LDSVEAENLTASINIEV	I	QVDVATNKQSSDIDET	TMFVI
	10	20	30	40	50	60
	70	80	90	100	110	120
gi-13621430.	EALDKESPLPNSVTT	TKGNGKTSFEQLTFSEVGQYHY	KIHQLLGKNSQYHYDET	VYEVV		
gi-13621430.	EALDKESPLPNSVTT	TKGNGKTSFEQLTFSEVGQYHY	KIHQLLGKNSQYHYDET	VYEVV		

FIGURE 59A

PCT/US06/27239

90 100 110 120

130 140 150 160 170 180

gi-13621430. IYVLYNEQSGALETNLVSNLGETEKSELIFKQEYSEKTPEPHQPDTEKEKPQKKRNGI
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||gi-13621430. IYVLYNEQSGALETNLVSNLGETEKSELIFKQEYSEKTPEPHQPDTEKEKPQKKRNGI
130 140 150 160 170 180

190 200 210

gi-13621430. LPSTGEMVSYVSALGIVLVATITLYSIYKKLTKSK

gi-13621430. LPSTGEMVSYVSALGIVLVATITLYSIYKKLTKSK
190 200 210

gi-13621430.pep

/home/morama/gas/pili/align/gi-19745305.pep

gi|19745305|ref|NP_606441.1| hypothetical protein [Streptococcus pyogenes MGAS82 32]

SCORES Init1: 163 Initn: 243 Opt: 273 z-score: 82.2 E(): 0.5

>>/home/morama/gas/pili/align/gi-19745305.pep (195 aa)

. initn: 243 init1: 163 opt: 273 z-score: 82.2 expect(): 0.5

Smith-Waterman score: 320; 31.9% identity in 213 aa overlap

(1-213:1-193)

10 20 30 40 50 60

gi-13621430. MKKSILRLAIYGILMSFCLLDSSVEAENLTASINIEVINQVDVATNKQSSDIDETFMFVI

gi-19745305. MRKYWKMLFSVVMILTMALFNQTVLAKDSTVQTSISVENVLERAGDSTS-----FSVAL

10 20 30 40 50

70 80 90 100 110 120

gi-13621430. EALDKESPLPNSVTTSVKGNKGTSFEQLTFSEVGQYHYKIHQLLGKNSQYHYDETUVYEVV

gi-19745305. ESIDAMKTI-DEIT--IAGSGKASFSPLETFTVQYTYRVAQPSQNQDYQADTTVFDVL

60 70 80 90 100 110

130 140 150 160 170 180

gi-13621430. IYVLYNEQSGALETNLVSNLGETEKSELIFKQEYSEKTPEPHQPDTEKEKPQKKRNGI

gi-19745305. VYVITYDED-GTLVAKVISRRAGDEEKSAITFFPKPKRLVPIPPIPQDPDI-----PKTP-----

120 130 140 150 160

190 200 210

gi-13621430. LPSTGEMVSYVSALGIVLVATITLYSIYKKLTKSK

gi-19745305. LPLAGEVKSLIGILSIVLLGLVLLYYV-KKLKSRL

170 180 190

gi-13621430.pep

/home/morama/gas/pili/align/gi-28810261.pep

gi|28810261|dbj|BAC63199.1| hypothetical protein [Streptococcus pyogenes SSI-1]

SCORES Init1: 164 Initn: 239 Opt: 268 z-score: 81.5 E(): 0.55

>>/home/morama/gas/pili/align/gi-28810261.pep (189 aa)

initn: 239 init1: 164 opt: 268 z-score: 81.5 expect(): 0.55

Smith-Waterman score: 306; 30.6% identity in 206 aa overlap

FIGURE 59B

205/487

(8-2132-187) US 505,272 39

gi-13621430.pep
/home/morrama/gas/pili/align/gi-19224139.pep

gi|19224139|gb|AAL86410.1|AF447492_7 unknown [Streptococcus pyogenes]

```
SCORES Init1: 164 Initn: 236 Opt: 265 z-score: 81.0 E(): 0.57  
>>/home/morama/gas/pili/align/gi-19224139.pep (189 aa)  
initn: 236 init1: 164 opt: 265 Z-score: 81.0 expect(): 0.57  
Smith-Waterman score: 303; 30.6% identity in 206 aa overlap  
(8-213:2-187)
```

	10	20	30	40	50	60
gi-13621430.	MKK	SILRILAIGYLLMSFC	LDSVEAENLTASINIEV	INQVDVATNKQSSDIDET	FMFVI	
gi-19224139.	MLFSVV	MLTMLAFNQT	VLA	DSTVQTSIS	VENVLERAGD	STP-----FSIAL
	10	20	30	40		
	70	80	90	100	110	120
gi-13621430.	EALDKESPL	PNSVTT	SVKGNGKT	SFEQLTFSEVG	OYHYKIHOLLG	KNSQYHYDET
gi-19224139.	ESIDAMKTEE	--ITIAGSGKA	FSPLTT	TVGQTYR	VYQKPSQN	KDYQADTTVFDVL
	50	60	70	80	90	100
	130	140	150	160	170	180
gi-13621430.	IYVLYNEQS	GALETN	LVSNLGE	TEKSEL	IFKQEYSE	KTPPEPHQPD
gi-19224139.	VYVTYDED	-GTLVAKV	ISRRAGDEEKS	KAIT	TFKPKRLV	KPIP
	110	120	130	140	150	
	190	200	210			
gi-13621430.	LPSTGEMVSY	V SALGIVL	VATITLYS	IYKKL	TKSK	
gi-19224139.	LPLAGEV	KSKLLGIL	SIVLLGL	VLLYV-	KKLKS	KL
	160	170	180			

FIGURE 59C

PCT/US05/27239

gi-13621430.pep

/home/morama/gas/pili/align/orf82.pep

TRANSLATE of: orf82.seq check: 4296 from: 1 to: 672
 generated symbols 1 to: 224.

GETSEQ from morama, September 13, 2004 17:09.

SCORES Init1: 163 Initn: 235 Opt: 264 z-score: 81.0 E(): 0.58
 >>/home/morama/gas/pili/align/orf82.pep (224 aa)
 initn: 235 init1: 163 opt: 264 z-score: 81.0 expect(): 0.58
 Smith-Waterman score: 304; 30.5% identity in 213 aa overlap
 (1-213:30-222)

gi-13621430.	10 20 30
	MKKSILRILAIGYLLMSFCLLDSVEAENLTA :: ::::: :: :::: :: ::
orf82.pep	LLFQRVKIFLLTIVLSVLFKNNEERRRLLRKYWKMLFSVVMILITMLAFNQTVLAKDSTV 10 20 30 40 50 60
gi-13621430.	40 50 60 70 80 90
orf82.pep	SINIEVINQVDVATNKQSSDIDETFMFVIEALDKESPLPNSVTTSVKGNGKTSFEQLTFS :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: QTSISVENLERAGDSTP-----FSVALESIDAMKTI-DEIT--IAGSGKASFSPLTFT 70 80 90 100 110
gi-13621430.	100 110 120 130 140 150
orf82.pep	EVGQYHYKIHQLLGKNSQYHYDETIVYEVVIYVLYNEQSGALETNLVSNKLGTEKESELIF :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: TVGQTYRYVYQKPSQNQDYQADTTVFDVLVYVTYDED-GTLVAKVISRRAGDEEKSATF 120 130 140 150 160 170
gi-13621430.	160 170 180 190 200 210
orf82.pep	KQEYSEKTPEPHQPDTEKEKPKQRNGILPSTGEMVSYVSAIGIVLVATITLYSIYKKL : :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: KPKRLVKPKIPPRQPNI-----PKTP---LPLAGEVKSLLGILSIVLLGLIVLLYV-KKL 180 190 200 210 220
gi-13621430.	KTSK
orf82.pep	: KSRL

gi-13621430.pep

/home/morama/gas/pili/align/gi-21909638.pep

gi|21909638|ref|NP_663906.1| hypothetical protein [Streptococcus pyogenes MGAS31
 5]

SCORES Init1: 164 Initn: 239 Opt: 261 z-score: 80.5 E(): 0.62
 >>/home/morama/gas/pili/align/gi-21909638.pep (183 aa)
 initn: 239 init1: 164 opt: 261 z-score: 80.5 expect(): 0.62
 Smith-Waterman score: 302; 31.5% identity in 200 aa overlap
 (14-213:2-181)

gi-13621430.	10 20 30 40 50 60
	MKKSILRILAIGYLLMSFCLLDSVEAENLTA :: ::::: :: ::: :: :: ::
gi-21909638.	MILTMLAFNQTVLAKDSTVQTSISVENLERAGDSTP-----FSIAL

PCT/US05/27239

10 20 30

40

70 80 90 100 110 120

gi-13621430. EALDKESPLPNSVTTSVKGNGKTSFEQLTFSEVGQYHYKIHQLLGKNSQYHYDETUVYEVV
|::| : |::| : |::| : |::| : |::| : |::| : |::| : |::| :gi-21909638. ESIDAMKTEEE---ITIAGSGKASFSPLTFTTVGQTYRVYQKPSQNKDQADTTVFDVL
50 60 70 80 90

130 140 150 160 170 180

gi-13621430. IYVLYNEQSGALETNLVSNKLGTEKSELIFKQEYSEKTPEPHQPDTEKEKPQKRNGI
|::| : |::| : |::| : |::| : |::| : |::| : |::| : |::| :gi-21909638. VVVTYDED-GTLVAKVISRRAGDEEKSAITFKPKWLVKPIPPRQPNI----PKTP----
100 110 120 130 140

190 200 210

gi-13621430. LPSTGEMVSYSALGIVLVATITLYSIYKKLTKS
|::| : |::| : |::| : |::| : |::| : |::| :gi-21909638. LPLAGEVKSLLGILSIVLLGLLVLLYV-KKLKSRL
150 160 170 180

gi-13621430.pep

/home/morama/gas/pili/align/gi-19745303.pep

gi|19745303|ref|NP_606439.1| hypothetical protein [Streptococcus pyogenes MGAS82
32]:

SCORES Init1: 121 Initn: 121 Opt: 126 z-score: 61.4 E(): 6.7
 >>/home/morama/gas/pili/align/gi-19745303.pep (344 aa)
 initn: 121 init1: 121 opt: 126 Z-score: 61.4 expect(): 6.7
 Smith-Waterman score: 126; 27.0% identity in 100 aa overlap
 (59-155:84-183)

30 40 50 60 70 80

gi-13621430. LTASTINIEVINQVDVATNKQSSDIDETFMFVIEALDKESPLP--NSVTTSVKGNG-KTSE
|::| : |::| : |::| : |::| : |::| : |::| :gi-19745303. LMPKADYTFKVEADDNAKGKTDGLDIKPGVIDGLENTKTIHYGNSDKTTAKEKSVNFDF
60 70 80 90 100 110

90 100 110 120 130 140

gi-13621430. EQLTFSEVGQYHYKIHQLLGKNSQYHYDETUVYEVVIVYVLYNEQSGALETNLVSNKLGTE
|::| : |::| : |::| : |::| : |::| : |::| : |::| : |::| :gi-19745303. ANVKFPGVGVYRYTVSEVNGNKAGIAYDSQQWTVDVYVNRDGGFEAKYIVSTEGGQSD
120 130 140 150 160 170

150 160 170 180 190 200

gi-13621430. KSELEIFKQEYSEKTPEPHQPDTEKEKPQKRNGILPSTGEMVSYSALGIVLVATITLY
|::| : |::| : |::| :gi-19745303. KKPVLFKNFFDTTSKVTKVTGNTGEHQRSFSFTLLLTPNECFEKQVNNILQGGETKK
180 190 200 210 220 230

gi-13621430.pep

/home/morama/gas/pili/align/gi-13621428.pep

gi|13621428|gb|AAK33238.1| hypothetical protein [Streptococcus pyogenes]

SCORES Init1: 58 Initn: 86 Opt: 122 z-score: 60.9 E(): 7.2
 >>/home/morama/gas/pili/align/gi-13621428.pep (340 aa)
 initn: 86 init1: 58 opt: 122 Z-score: 60.9 expect(): 7.2
 Smith-Waterman score: 135; 29.1% identity in 172 aa overlap

(8-159E6#17D) / US05 / 27239

	10	20	30	40	50
gi-13621430.	MKKSTILRILAIGYLLMSFCLLDSVEAENLTASINIEVINQVD-VATNKQSSDIDETFMF-	: : : : : : : : : : : : :			
gi-13621428.	MKLRHILLTGAALTSFAAT-TVHGETVNGAKLTVKNLDLVNSNALIPNTDFTFKIE	10 20 30 40 50			

	60	70	80	90	100
gi-13621430.	VIEALDK-----ESPLPN-SVTTSVKGNGKTSFEQLTFSEV-----GQYHYKL-H	: :			
gi-13621428.	PDTITVNEGDGNKFKGVALNTPMTKVYTNTNSDKGGSNKTAEFDFSEVTFEKPGVYYYKVTE	60 70 80 90 100 110			

	110	120	130	140	150	160
gi-13621430.	QLLGKNSQYHYDETIVYEVVIVLYNE-QSGALETNLVSNLGETEKSELIFKQEYSEKTP	: : : : : : : : : : : : : : : : : : :				
gi-13621428.	EKIDKVPGVSYDFTTSYTVQVHLWNEEQQKPVATYIVGYKEGS--KVPIQFKNSLDSTTL	120 130 140 150 160 170				

	170	180	190	200	210
gi-13621430.	EPHQPDTEKEKPQKKRNGILPSTGEMVSYSVSLGIVLVATITLYSIYKKLKTSK	180 190 200 210 220 230			
gi-13621428.	TVKKVSGTGGDRSKDFNFGILTAKQYYKASEKVMILEKTTKGQAPVQTEASIDQLYHF				

gi-13621430.pep

/home/morrama/gas/pili/align/gi-19224137.pep

gi|19224137|gb|AAL86408.1|AF447492_5 EftLSL.A [Streptococcus pyogenes]

SCORES Init1: 88 InitN: 88 Opt: 119 z-score: 60.4 E(): 7.5
 >>/home/morrama/gas/pili/align/gi-19224137.pep (342 aa)
 initn: 88 init1: 88 opt: 119 Z-score: 60.4 expect(): 7.5
 Smith-Waterman score: 119; 29.7% identity in 111 aa overlap
 (72-176:93-201)

	50	60	70	80	90
gi-13621430.	DVATNKQSSDIDETFMFVIEALDKESPLPNVTTSVKGNGK-----TSFEQLTFSEVGQY	: :			
gi-19224137.	SVNPDSAATGTESNLPIKPGTAVNNQDIKVSVSNTDKTSGKEKQVVVDFMKVTFPSVGIY	70 80 90 100 110 120			

	100	110	120	130	140	150
gi-13621430.	HYKIHQLLGKNSQYHYDETIVYEVVIVLYNE-QSGALETNLVSNLGETEKSELIFKQEY	: :				
gi-19224137.	RYVVTEKGTAEGVTYDDTKWLVBVYVGNNEKGG-LEPKYIVSKKGDSATKEPIQFNNSF	130 140 150 160 170 180				

	160	170	180	190	200	210
gi-13621430.	SEKTPPEPHQPDTEKEKPQKKRNGILPSTGEMVSYSVSLGIVLVATITLYSIYKKLKTSK	: :				
gi-19224137.	-ETTSLKIEKEVTGNTGDHKKAFTFTLTLQPNEYYEASSVVKIEENGQTDVDKIGEAYKF	190 200 210 220 230 240				

gi-13621430.pep

/home/morrama/gas/pili/align/gi-50913503.pep

gi|50913503|ref|YP_059475.1| Fibronectin-binding protein [Streptococcus pyogenes MGAS10394]

PCT/US05/27239

SCORES Init1: 73 Initn: 73 Opt: 117 z-score: 60.4 E(): 7.6
 >>/home/morama/gas/pili/align/gi-50913503.pep (627 aa)
 initn: 73 init1: 73 opt: 117 z-score: 60.4 expect(): 7.6
 Smith-Waterman score: 118; 28.7% identity in 87 aa overlap
 (129-215:549-625)

100	110	120	130	140	150
gi-13621430.	KIHQLLGKNSQYHYDET	VYEVVIVLYNEQSGALET	NLVSNKLGETEKSELIFKQEYSEK		
			: : : :		
520	530	540	550	560	570
gi-50913503.	IETEDTKEPEVLMGGQSESVEFTKD	TQMSGFSETATVV	---EDTRPKLVFHFDNNEP		
160	170	180	190	200	210
gi-13621430.	TPEPHQPDTEKEKPQKKRNGILPSTGEMVS	YVSVSALGIVLVATITLYSIYKKLKTSK			
	: : : : : :	: : : : :			
580	590	600	610	620	
gi-50913503.	KVEEN	REKPTKNITPILPATGDIENVLAFLGILILSVLSIFSLLKNQSNKKV			

gi-13621430.pep

/home/morama/gas/pili/align/gi-19224134.pep

gi|19224134|gb|AAL86405.1|AF447492_2 protein F [Streptococcus pyogenes]

SCORES Init1: 73 Initn: 73 Opt: 115 z-score: 60.1 E(): 7.8
 >>/home/morama/gas/pili/align/gi-19224134.pep (698 aa)
 initn: 73 init1: 73 opt: 115 z-score: 60.1 expect(): 7.8
 Smith-Waterman score: 115; 27.4% identity in 73 aa overlap
 (143-215:631-697)

120	130	140	150	160	170
gi-13621430.	DET	VYEVVIVLYNEQSGALET	NLVSNKLGETEKSELIFKQEYSEK	TPEPHQPDTEKEK	
			: : : :		
610	620	630	640	650	
gi-19224134.	VLMGGQSESVEFTKD	TQMSGFSETVTIVEDTRPKLVFHFDNNEP	KVEEN	REK	
180	190	200	210		
gi-13621430.	PQKKRNGILPSTGEMVS	YVSVSALGIVLVATITLYSIYKKLKTSK			
	: : : : :	: : : :			
660	670	680	690		
gi-19224134.	PTKNITPILPATGDIENVLAFLGILILSVLSIFSLLKNQNNKV				

! Distributed over 1 thread.

! Start time: Wed Sep 15 18:45:36 2004

! Completion time: Wed Sep 15 18:45:38 2004

! CPU time used:

! Database scan: 0:00:00.1

! Post-scan processing: 0:00:00.3

! Total CPU time: 0:00:00.4

! Output File: gi-13621430.fasta

TYPE 3 pilus motifs / US605/27239
 protein F2 like fibronectin-binding protein
 Length: 696-733
 LPXTG
 pilin motif consensus PK (X₇) K
 E box consensus ETxAPxGY

SpyM3_0104/21909640	pilin motif	155 PKEKPIIYFK 398
	E box	YTFVETAAPDGY
SPS0106/28895018	pilin motif	269 PKEKPIIYFK 512
	E box	YTFVETAAPDGY
SpyM18_0132/19745307	pilin motif	269 PKEKPIIYFK 512
	E box	YTFVETAAPDGY
orf84	pilin motif	269 PKEKPIIYFK 512
	E box	YTFVETAAPDGY

TYPE 4 pilus motifs
 protein F2 like fibronectin-binding protein
 Length: 1161
 LPXTG
 pilin motif consensus PK (X₇₋₈) K
 E box consensus YxLxETxAPxGY

The protein is longer than the proteins belonging to TYPE 3 and has 4 possible pilin motifs and 2 E boxes

19224141	pilin motifs	215 PKGISQDIPK 571 PKGYQQVTEK 156 PKMSVVSKYGK 674 PKYDAKNQEYK
	E boxes	563 YDLYETKAPKGY 940 YTFVETAAPDGY

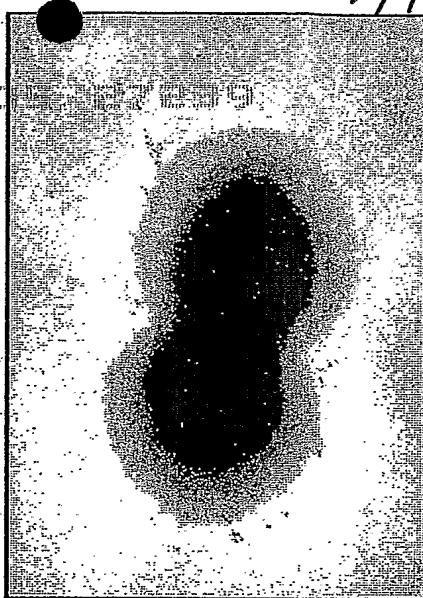
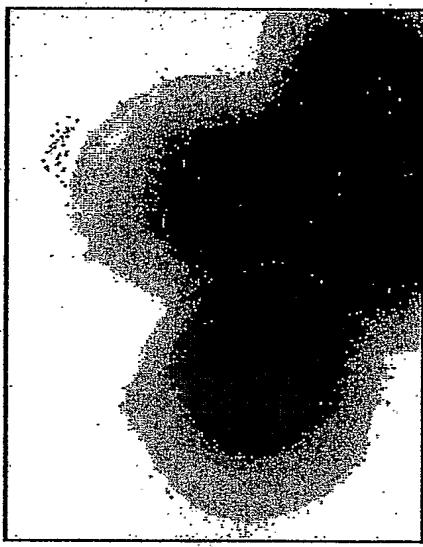
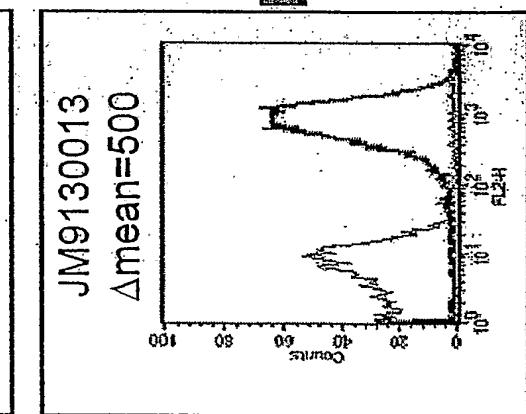
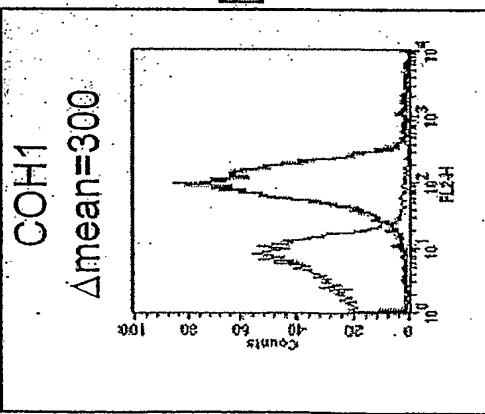
Formation of pili structures on GBS appears to be correlated to FACS values for surface expression of GBS80 protein

WO 2006/078318
PCT/US05/027239

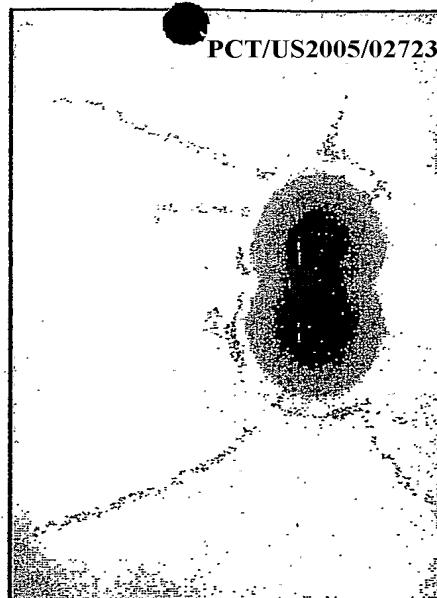
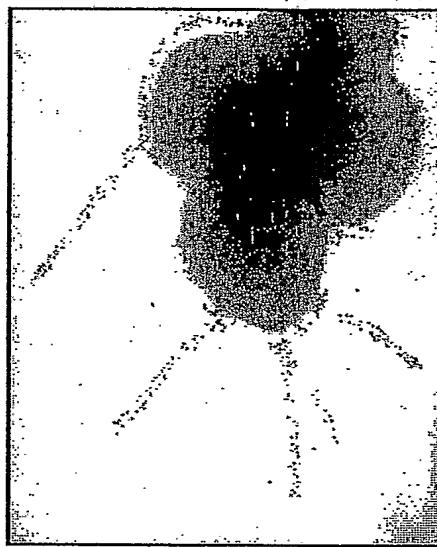
21/1/87

Immunogold Electron Microscopy

α -GBS80



α -GBS80



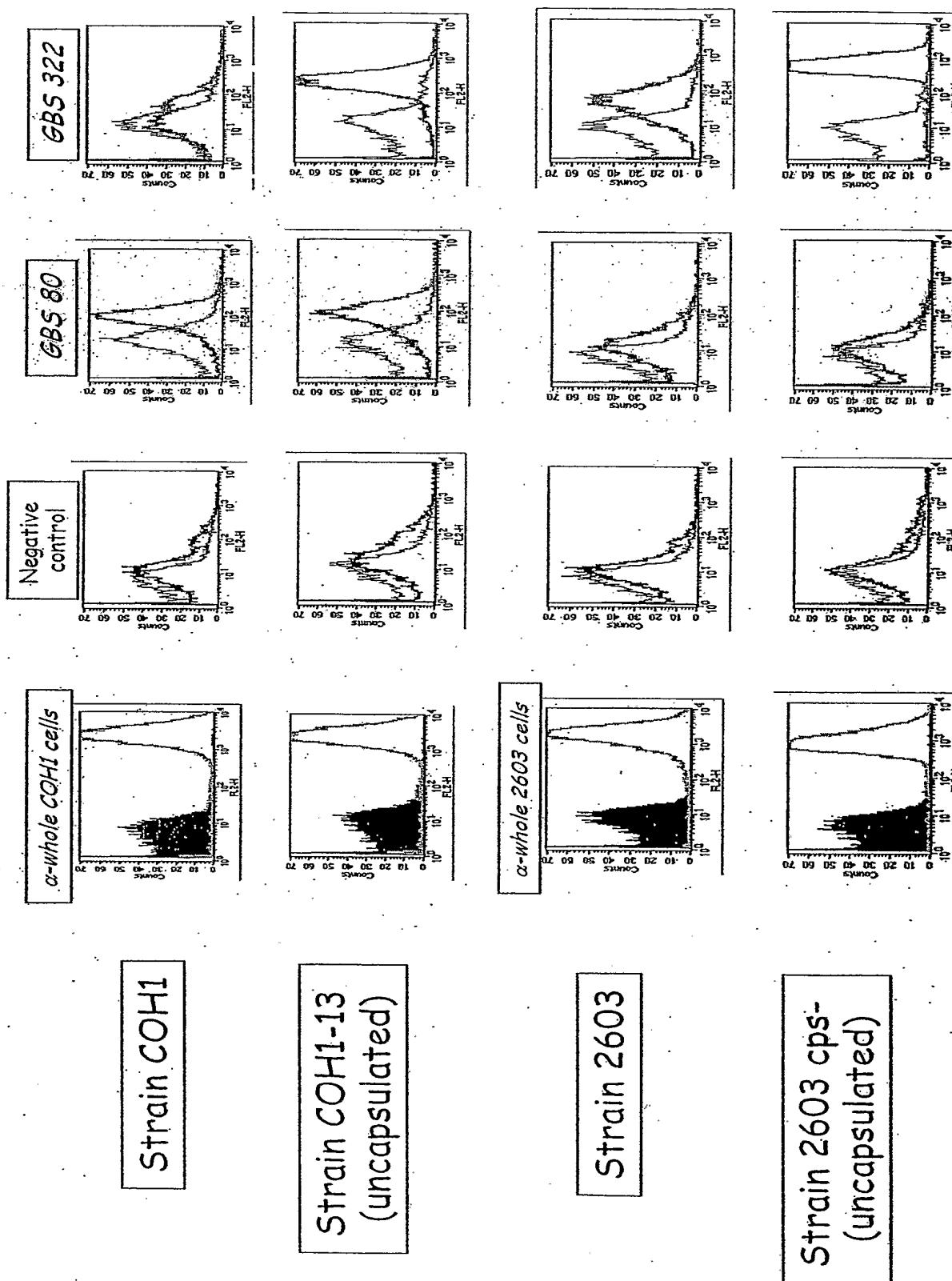
PCT/US2005/027239

CHIRON | VACCINES

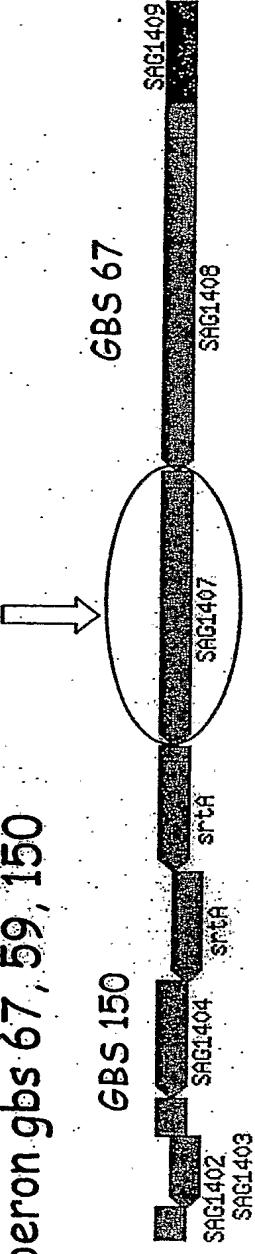
2121487

Surface exposure is capsule-dependent
for GBS 322 but not for GBS 80

FIGURE 62



**Adhesin island 2-
Operon gbs 67, 59, 150**



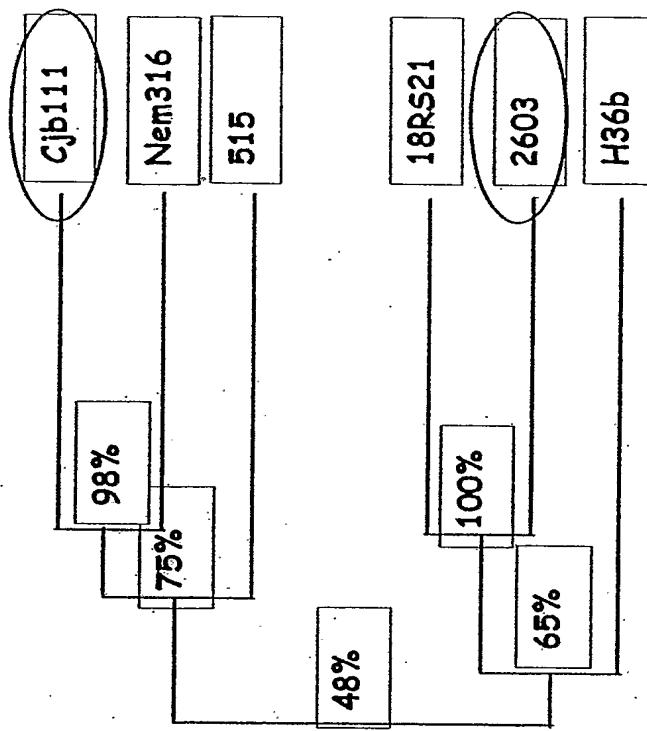
• 30 % identity with GBS 80

• By comparing GBS 59 amino acidic sequence of 2603 with that of other sequenced genomes, the following homologies are obtained:

2603	47%
nem316	62%
h36b	48%
515	48%
cjb111	100%
18rs21	not present (Sppb1)
coh1	not present (Sppb1)
A909	

• CGH: 1/20 GBS strains analyzed (18RS21)

Two-by-two amino acidic sequence comparison

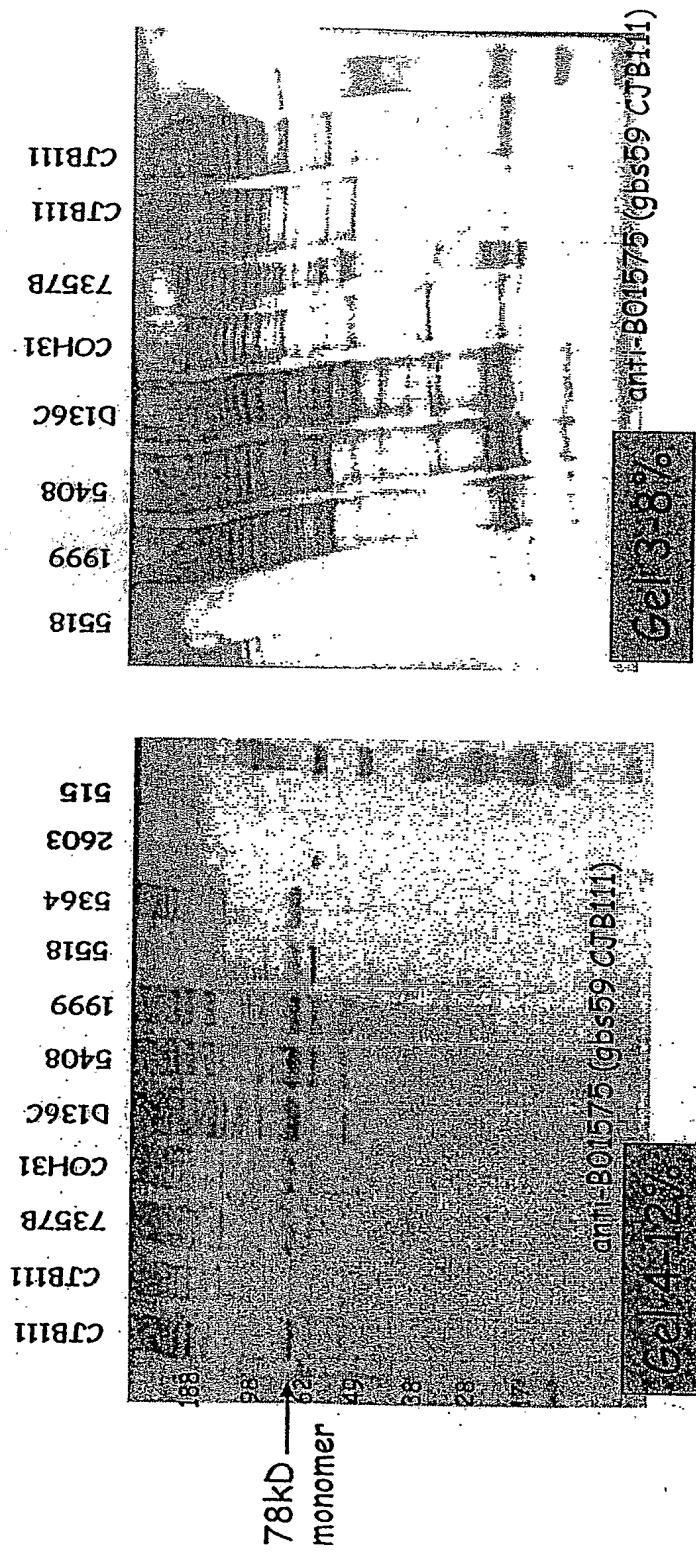


There seem to be two clusters suggesting the presence of two major isoforms

FIGURE 63

214/487

Western blotting with whole extracts derived from GBS strains



GBS 59 is part of a high molecular weight polymer (pilus) in GBS strains: CJB111, 7357b, coh31, d1363c, 5408, 1999, 5364, 5518, 515

Western blotting with purified proteins and whole extracts derived from GBS strains

WO 2006/078318

PCT/US05/27239

215/487

PCT/US2005/027239

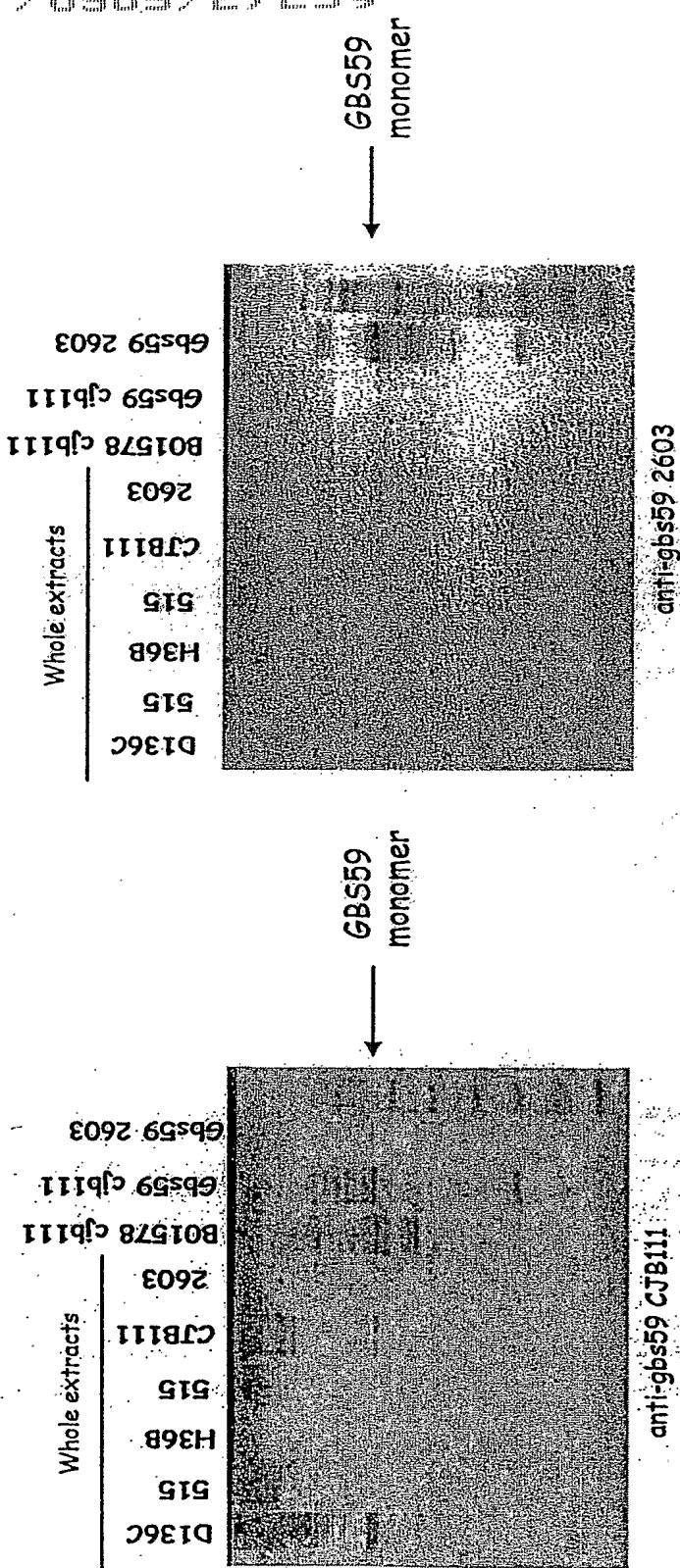


FIGURE 65

FACS analysis using mouse antiserum after immunization with BO1575 (gb559) from CJB111 genome

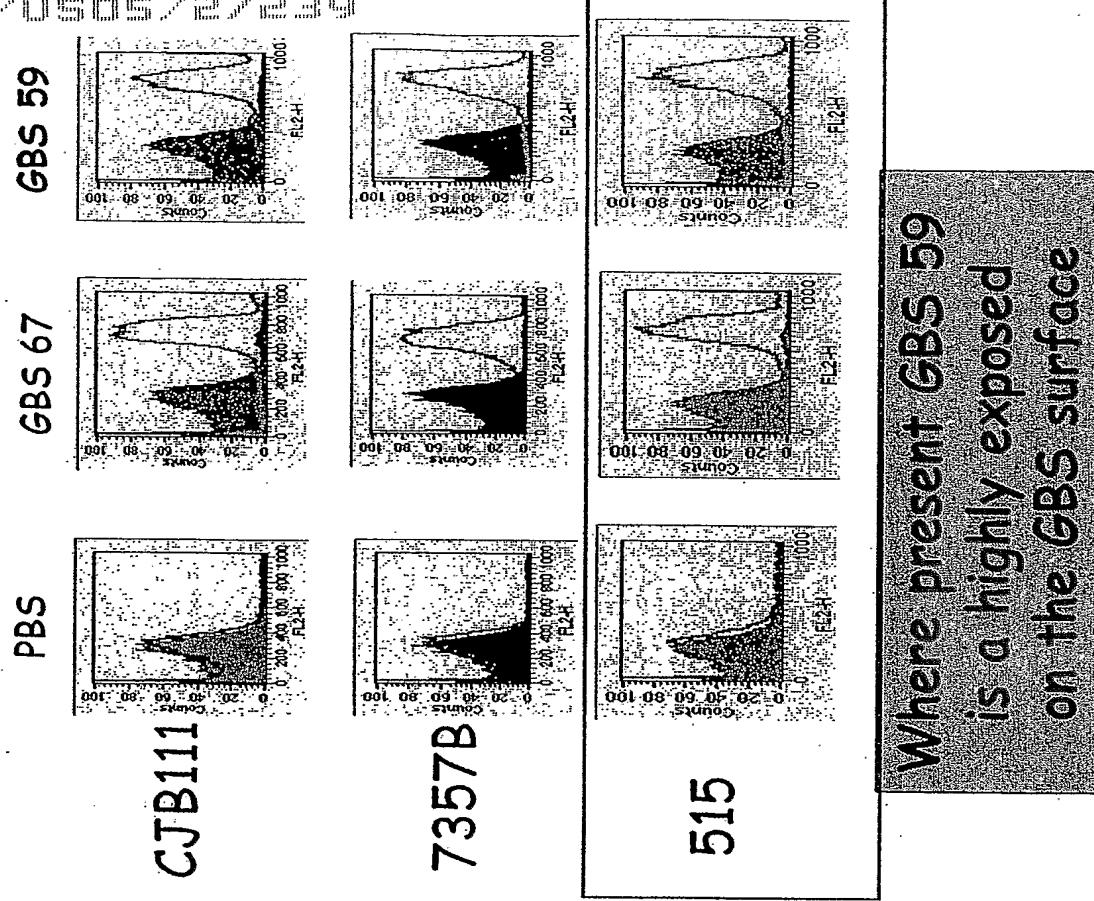
WO 2006/078318

216/487

PCT/US2005/027239

PCT/US05/27239

GBS strains	Type	FACS (Δ Mean)
DK1		565
DK8		559
Davis	Ia	577
5115		583
2986		443
5551		524
7357b-		596
5518		190
D136C		504
C0H31	III	505
DK21	II	249
CJB111		493
5364	V	593
2110		590
1999		594
2210		536
5408		567
1169		227



Where present GBS 59
is a highly exposed
on the GBS surface

FIGURE 66

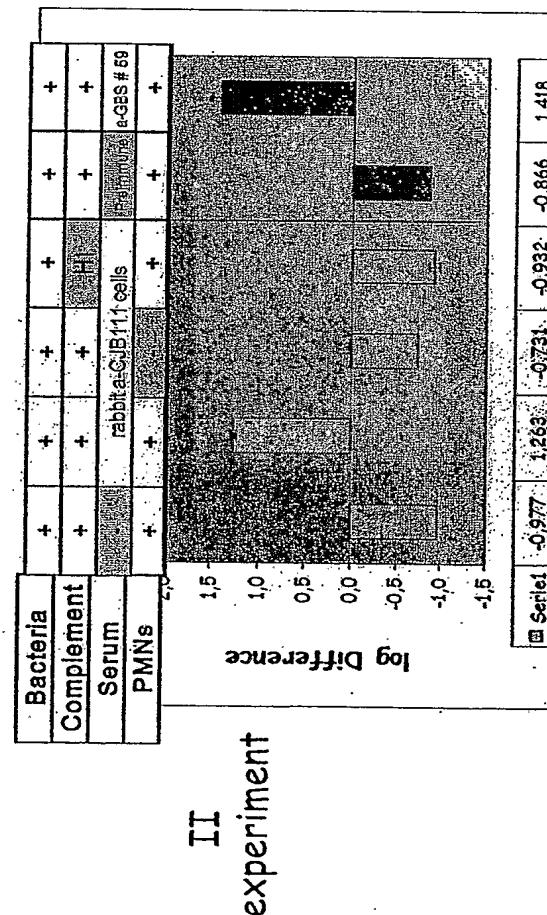
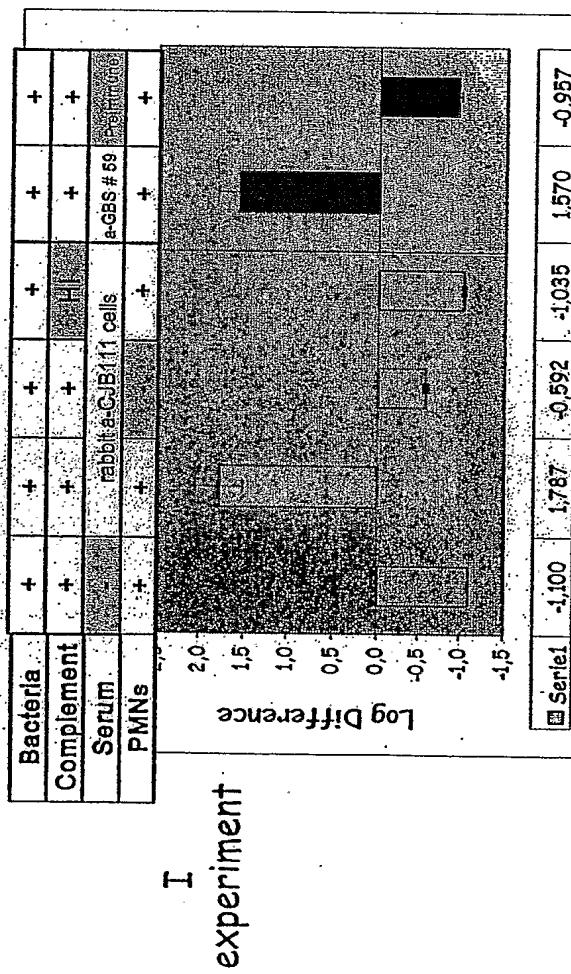
Opsonophagocytosis assays: B01575 (gbss59-cjb111)

WO 2006/078318

PCT/US05/27239

217/487

PCT/US2005/027239



• CJB111 GBS strain type Ia
 • Baby rabbit complement
 • Human PMNs
 • Positive control: anti-type V cells
 (rabbit serum anti fixed type V cells)

FIGURE 67

Association GBS 80-104

WB a-80, a-104 JM9130013 Total Extract

WO 2006/078318

PCT/US05/27239

218/487

PCT/US2005/027239

FACS	(ΔMEAN)
GBS 80	597
GBS 104	446

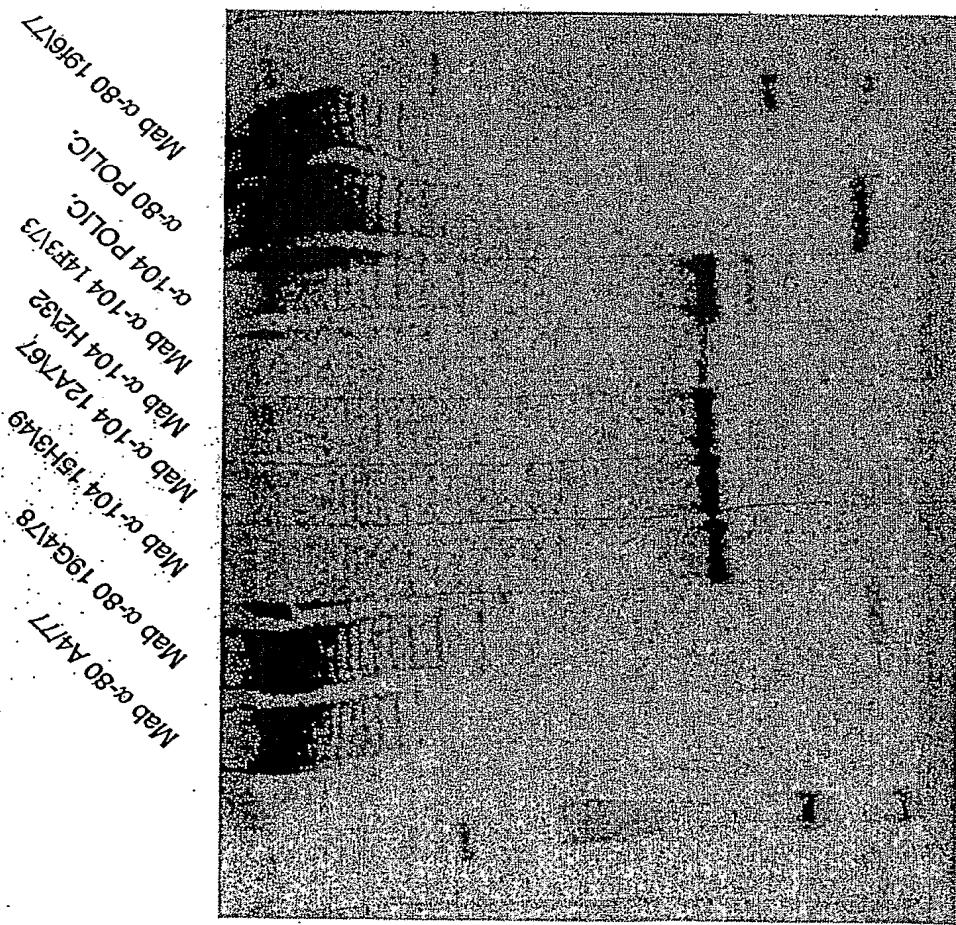
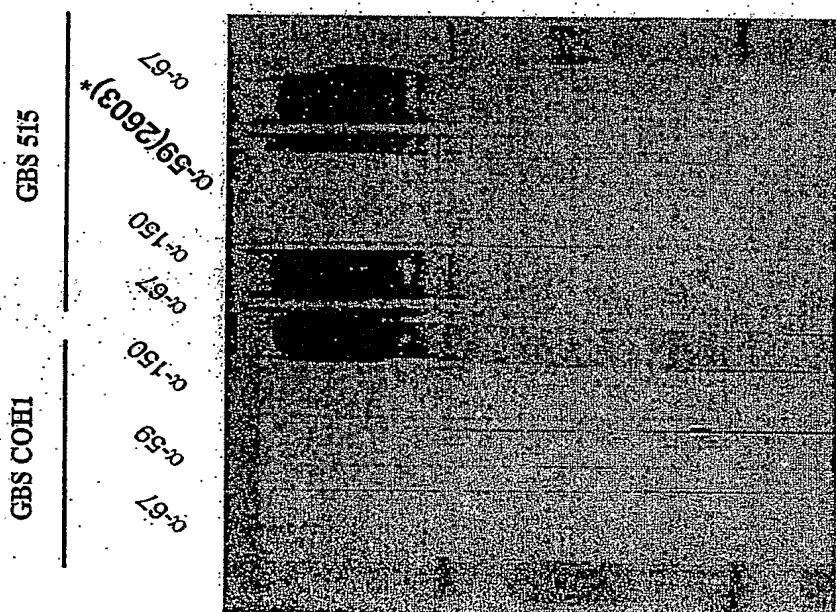


FIGURE 68

PCT/US05/27239

**WB GBS 515 Total Extract
 α -67; α -150; α -59 (2603)**



Controls:
GBS COH1 total extracts

- anti-gbs59 mouse serum after immunization with SAG1407 (GBS 59) from 2603 genome

GBS 67 and GBS 150 are parts of a high molecular weight polymer (pilus) in 515 GBS strain

Western Blotting ko GBS67
from 515 genome (clone 1.45)

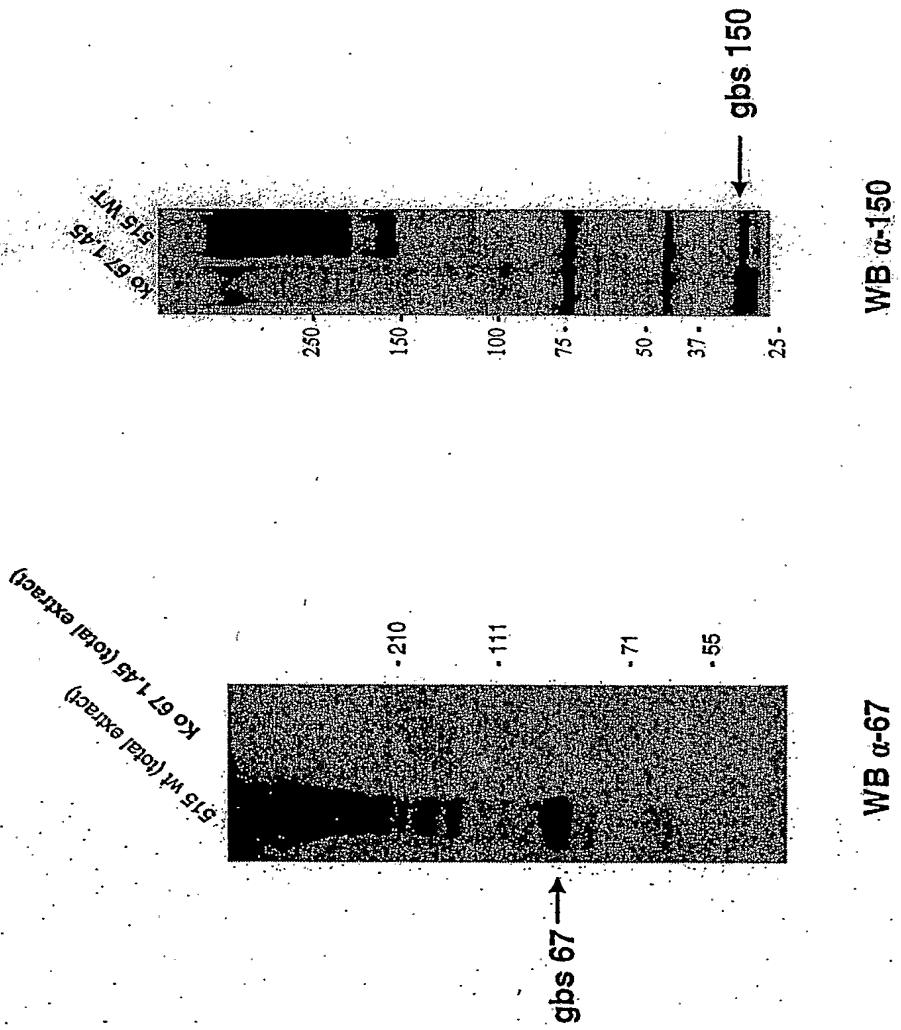


FIGURE 70

WO 2006/078318

220/487

PCT/US2005/027239

PCT/US05/27239

FACS GBS 515 Δ67

WO 2006/078318

PCT/US05/27239

221/487

PCT/US2005/027239

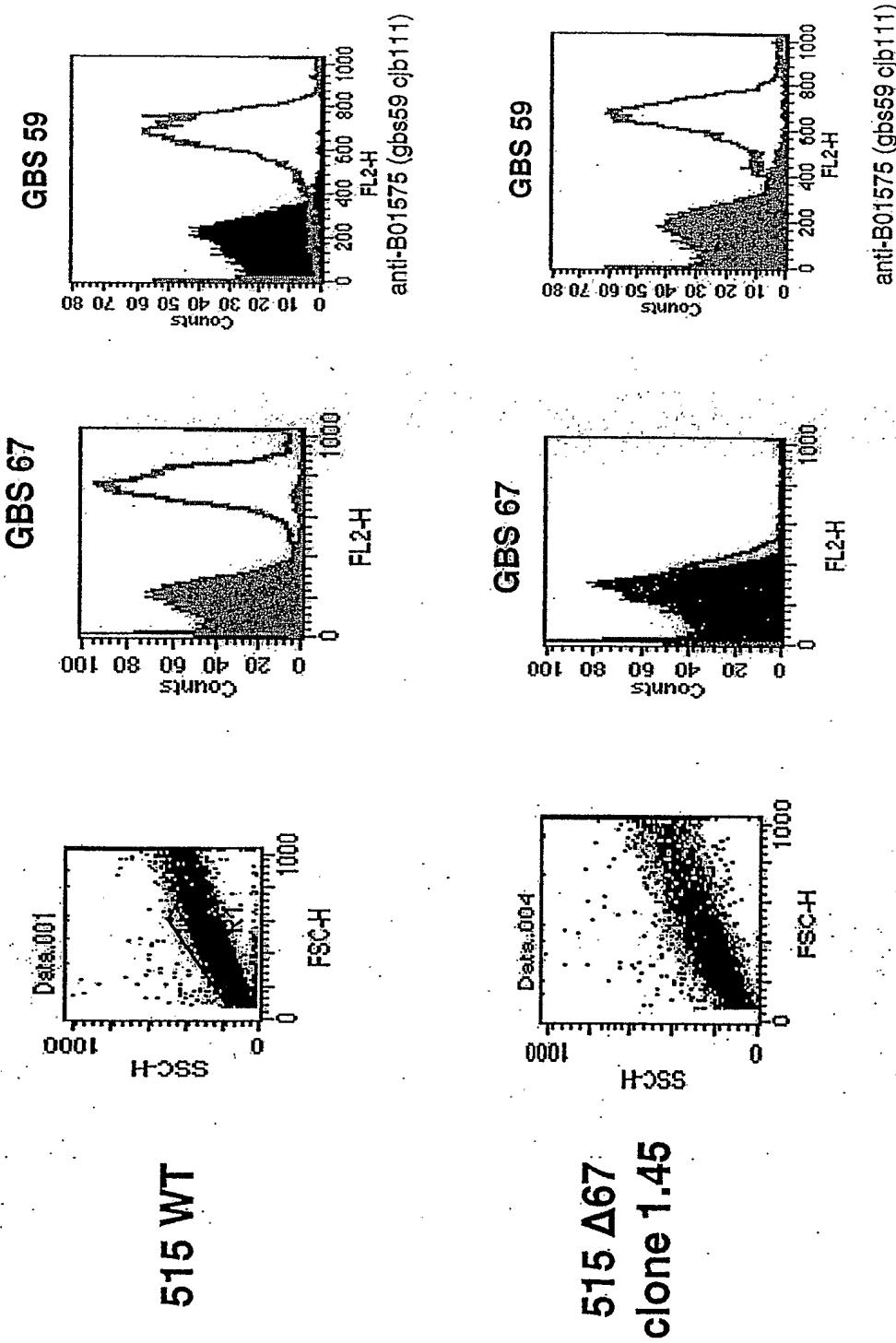


FIGURE 71

Complementation of GBS 515 KO 67 with pAM401-gbs80

WO 2006/078318

PCT/US05/27239

222/487

PCT/US2005/027239

GBS 80 forms a high molecular weight complex
(pilus) in absence of GBS 67

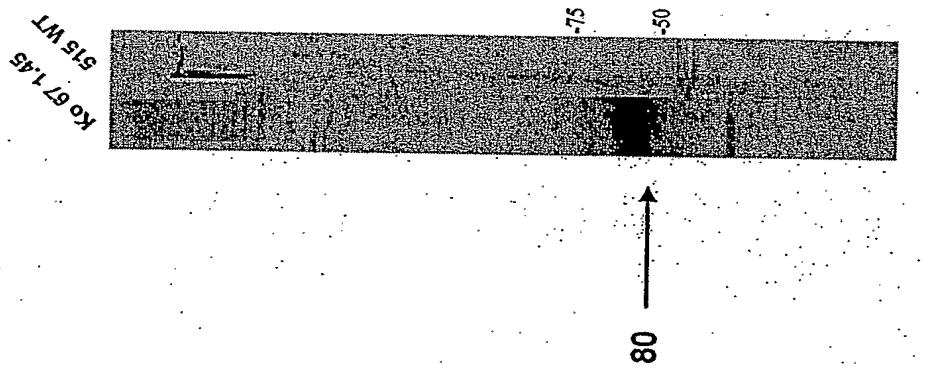


FIGURE 72

spyM6_0159 type I pilus present in M6

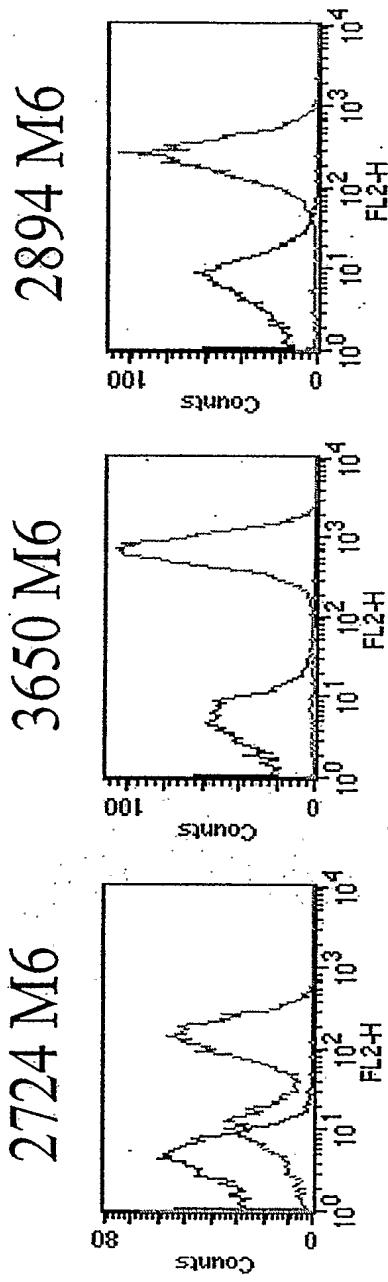


Figure 73

spyM6_0160 type 1 pilus present in M6

WO 2006/078318

PCT/US05/27239 224/487

PCT/US2005/027239

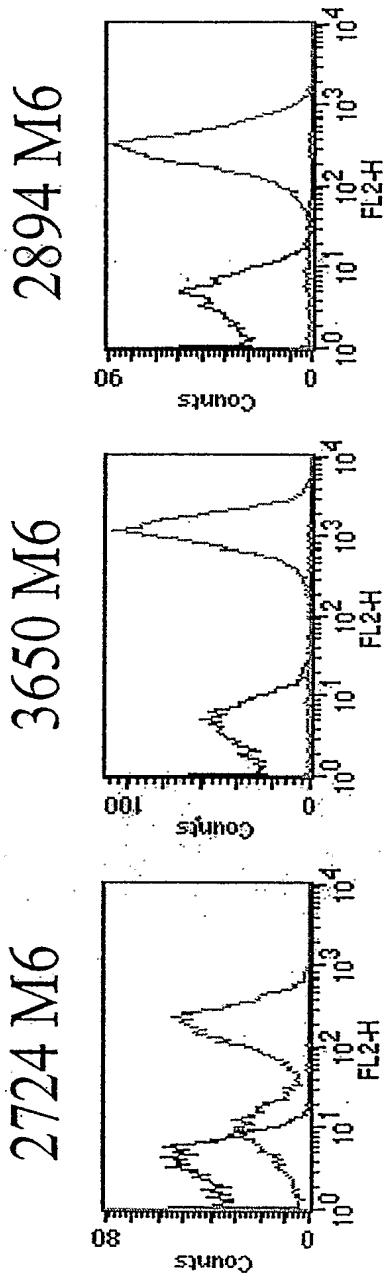


Figure 74

225/487

Gas15 type 2 pilus present in M1

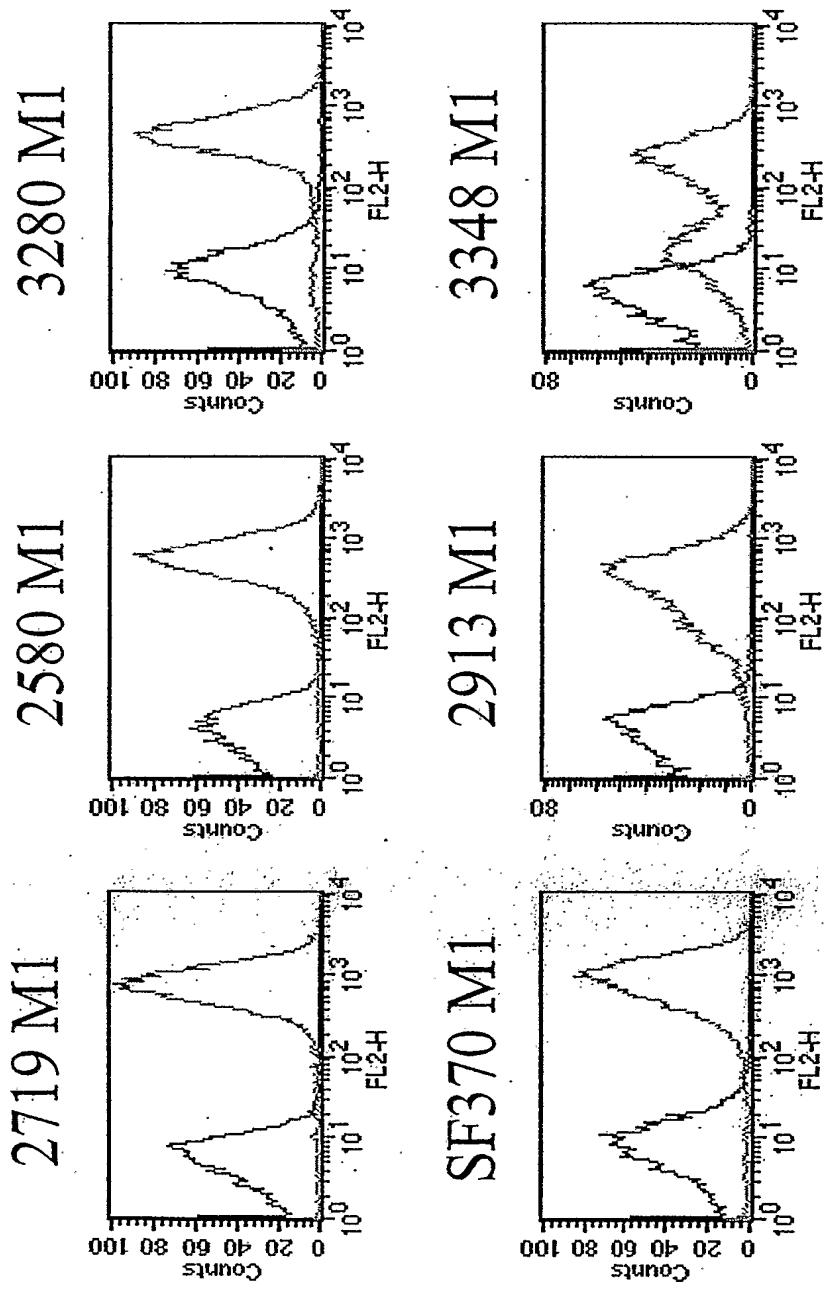


Figure 75

226/487

Gas16 type 2 pilus present in M1

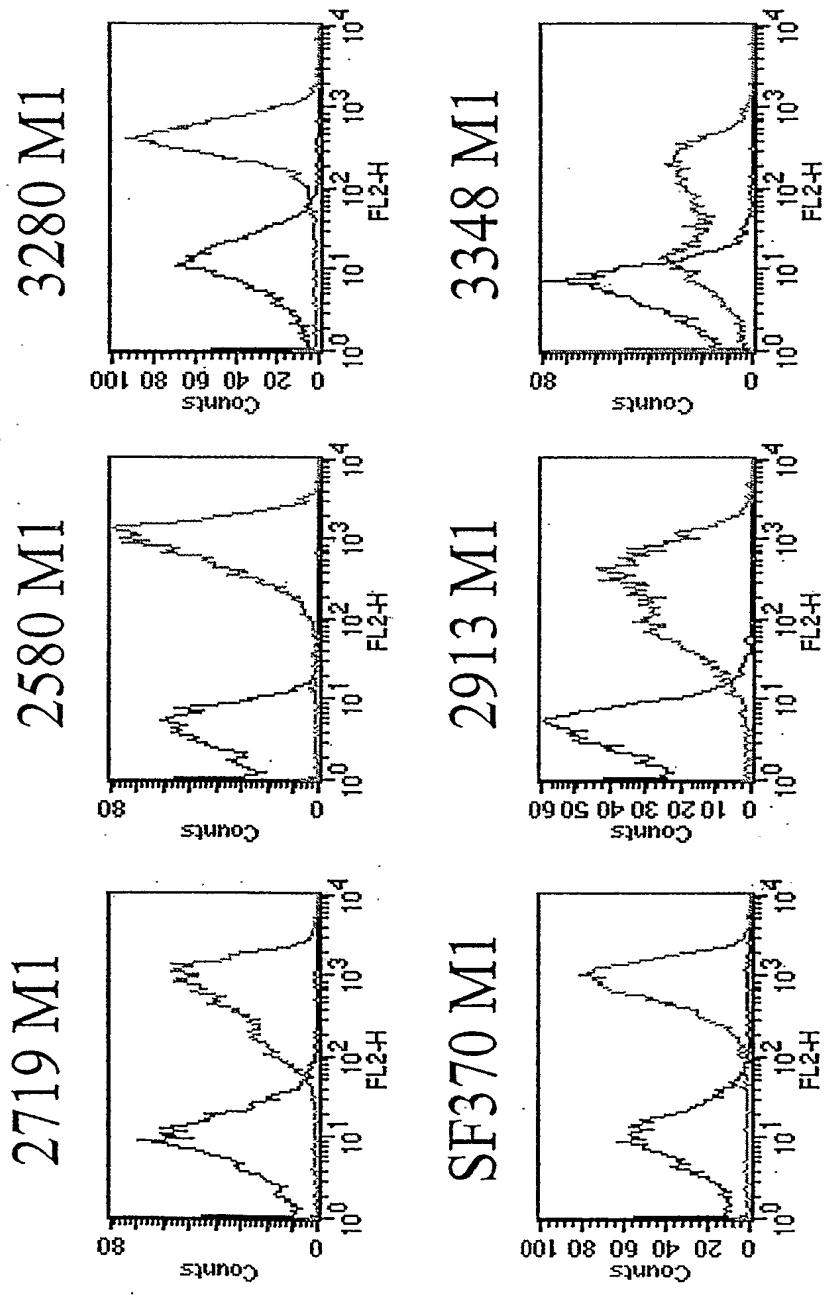


Figure 76

227/487

Gas18 serum 1 type 2 pilus present in M1

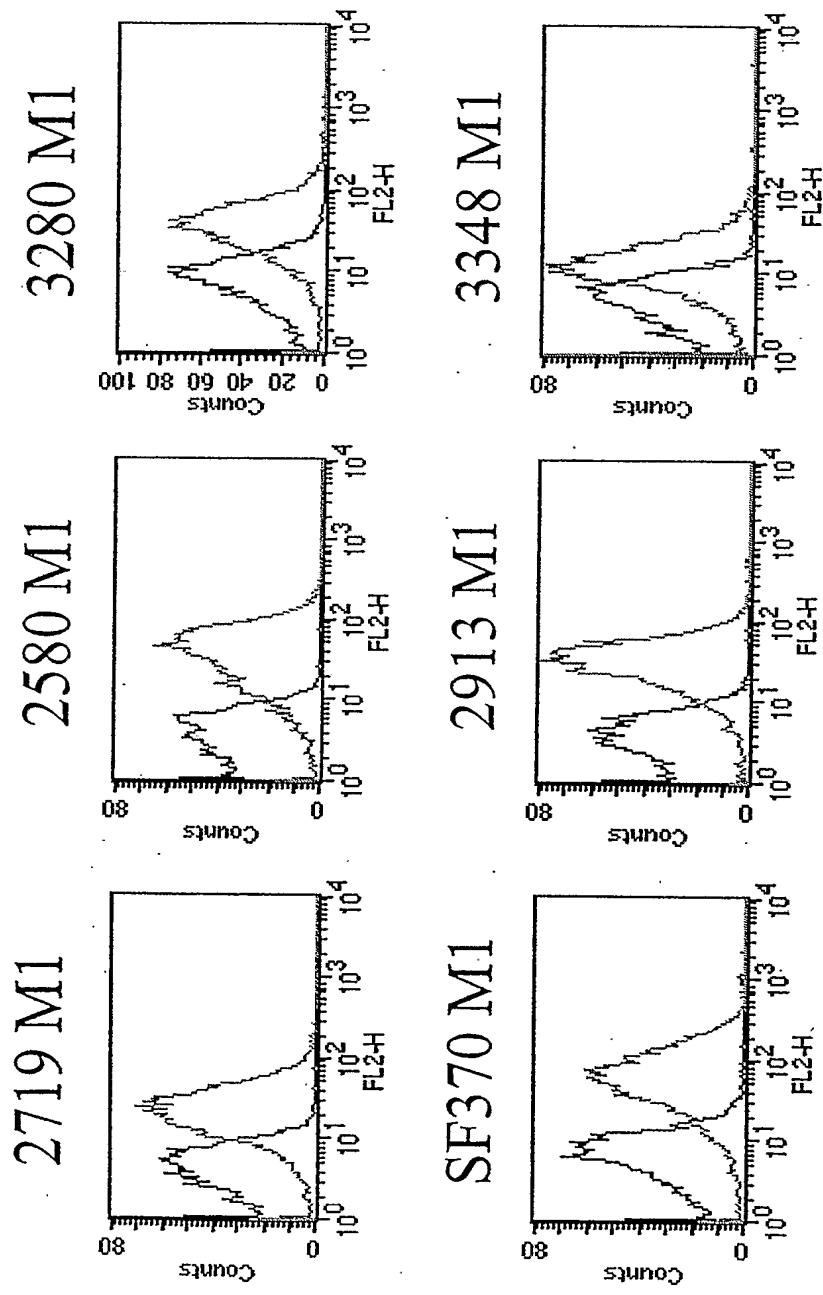


Figure 77

228/487

Gas18 serum 2 type 2 pilus present in M1

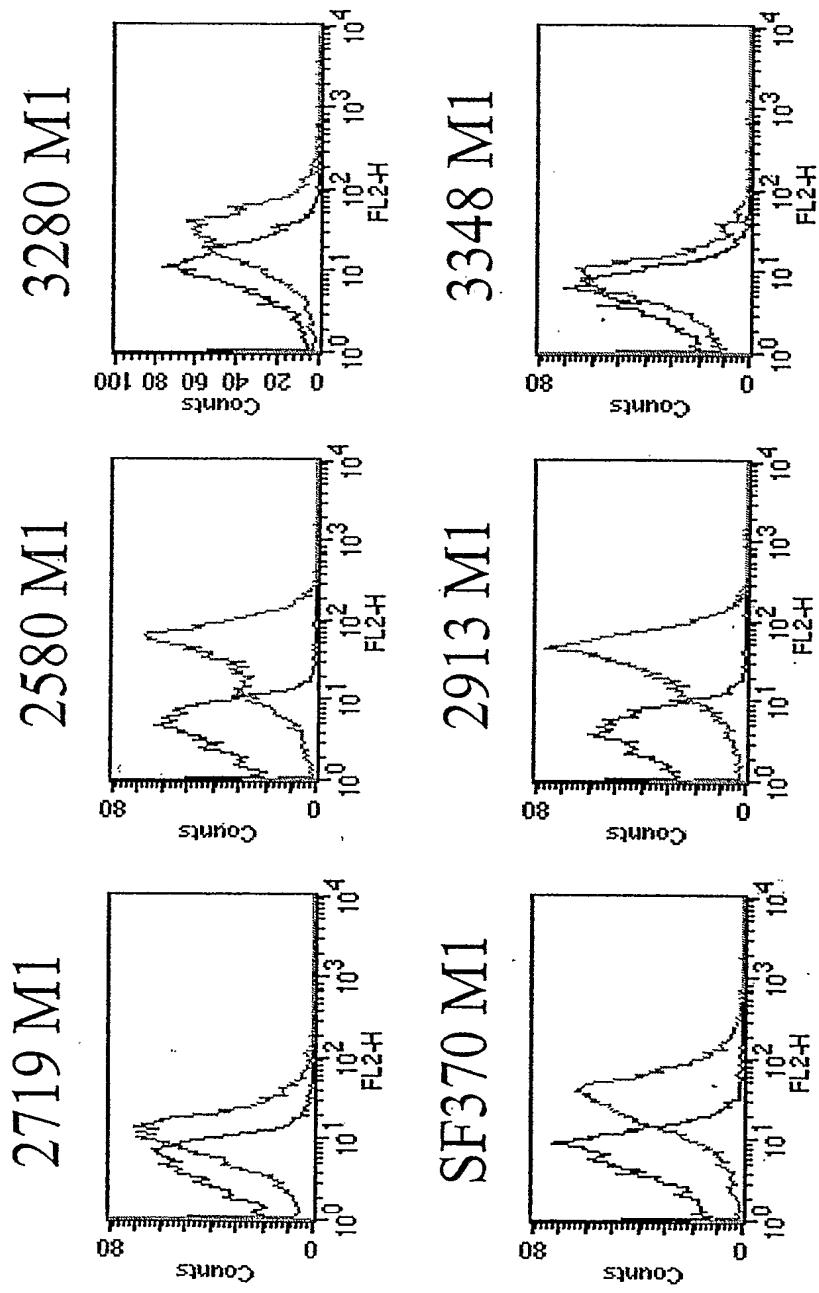


Figure 78

Gas16p2 type 2 pilus present in M1

WO 2006/078318

PCT/US2005/27239

PCT/US2005/027239

229/487

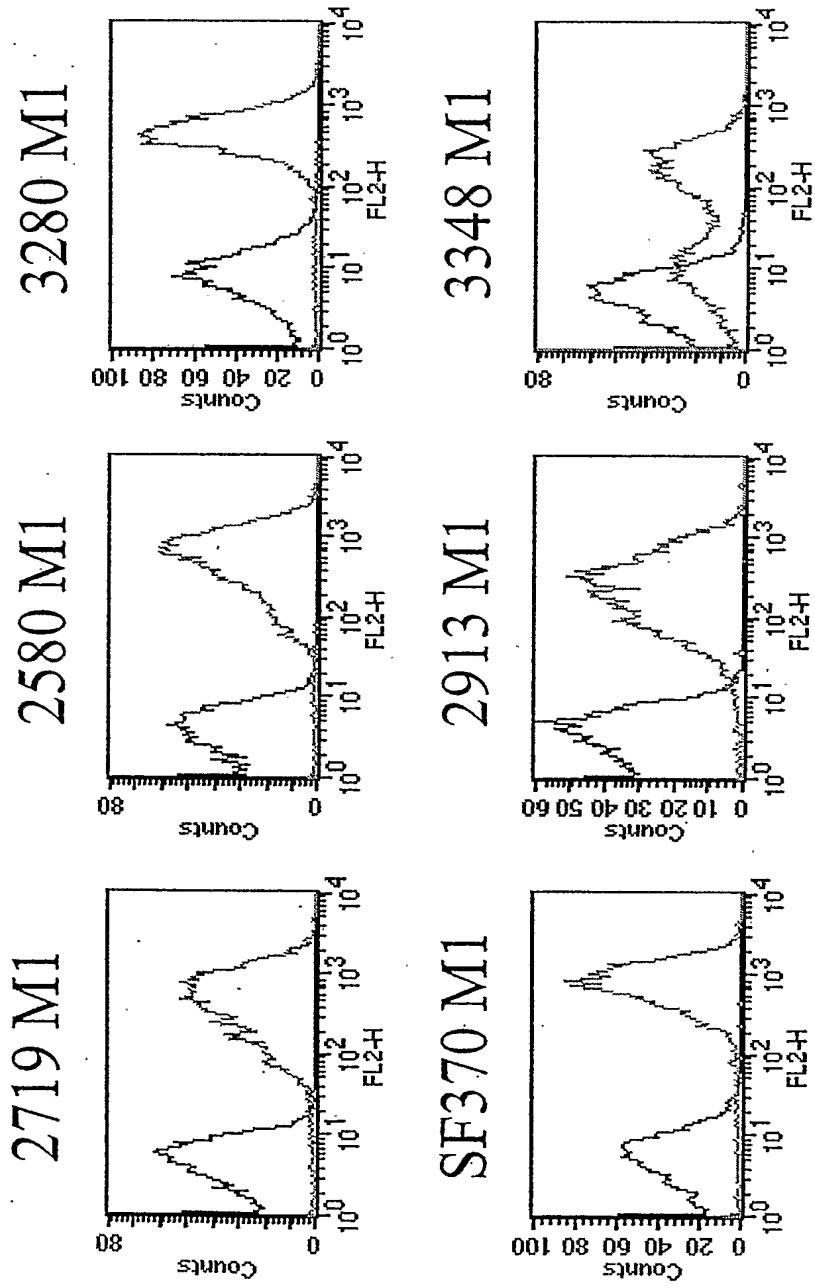


Figure 79

230/487

spyM3_0098 type 3 pilus present in M3

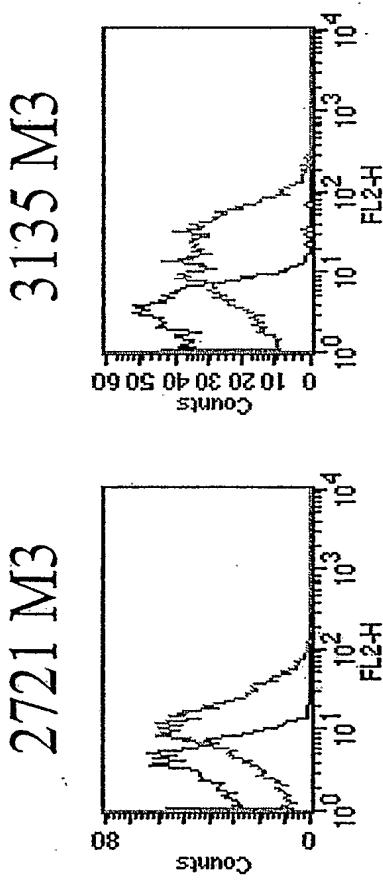
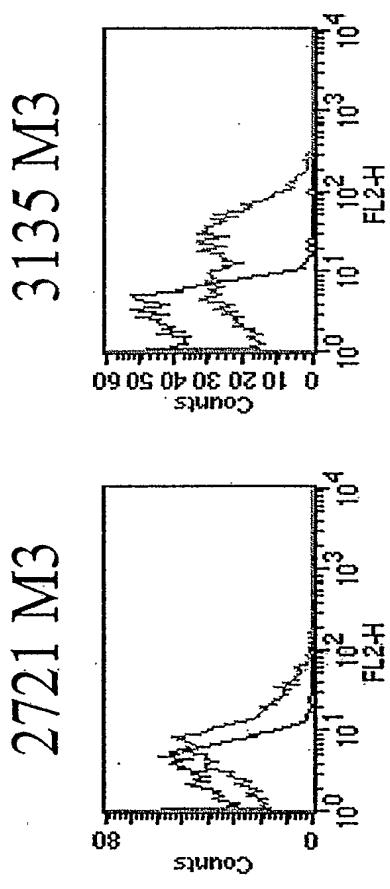


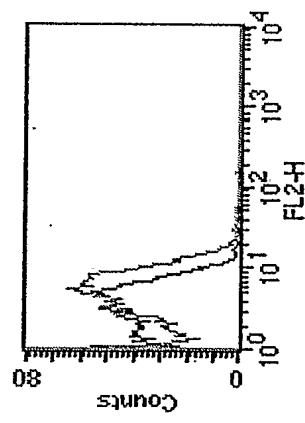
Figure 80

231/487

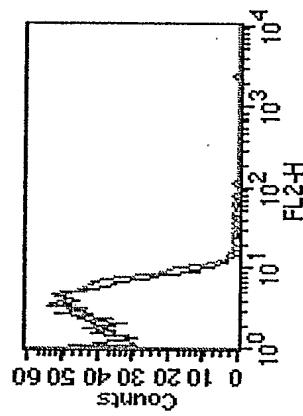
spyM3_0100 type 3 pilus present in M3**Figure 81**

spyM3_0102 type 3 pilus present in M3

2721 M3



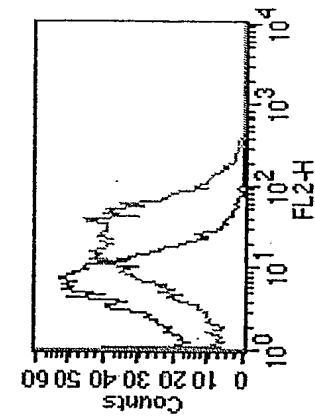
3135 M3



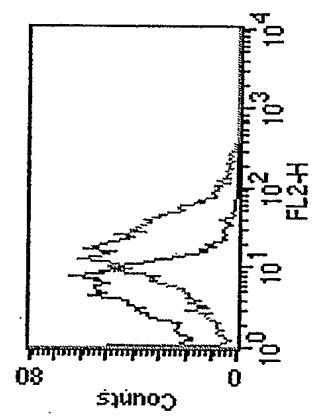
232/487

PCT/US2005/027239

2894 M6



3650 M6



2724 M6

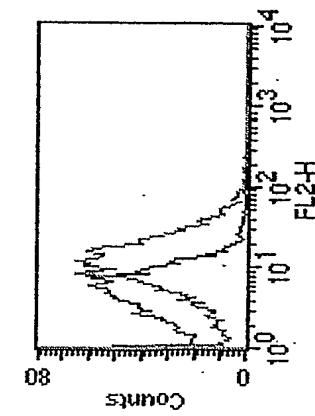
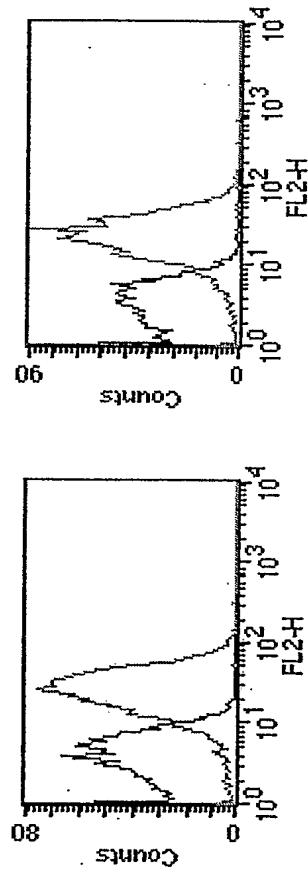
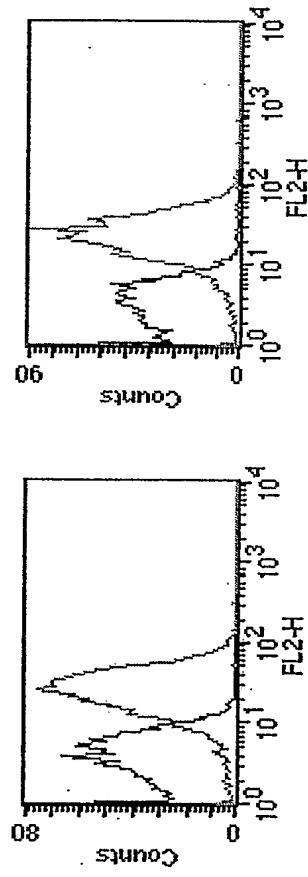
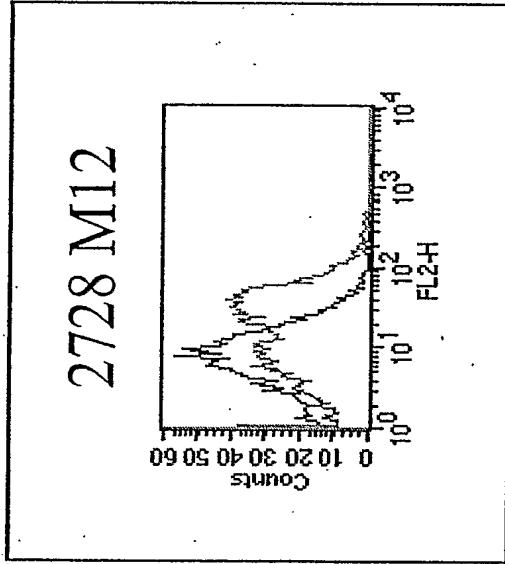


Figure 82

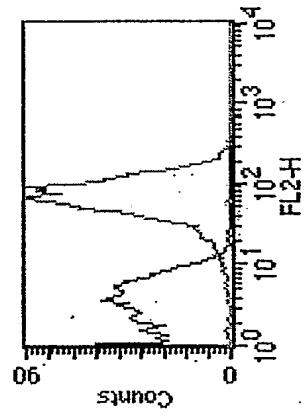
233/487
PCT/US2005/27239**spyM3_0104 type 3 pilus present in M3****2721 M3****3135 M3****2728 M12****Figure 83**

PCT/US05/27239

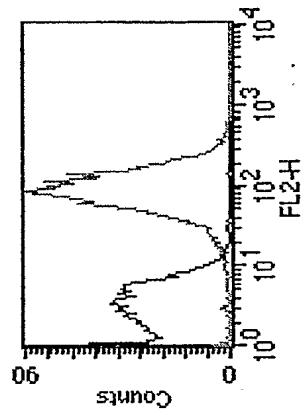
234/487

spyM3_0106 type 3 pilus present in M3

2721 M3



3135 M3



2728 M12

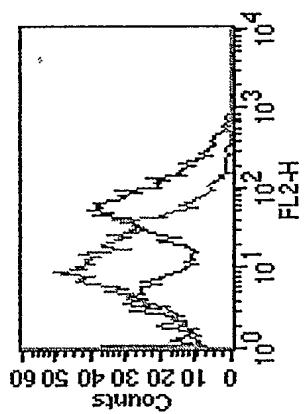
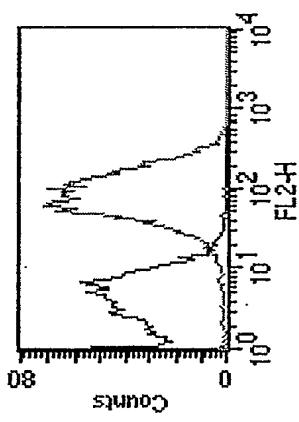


Figure 84

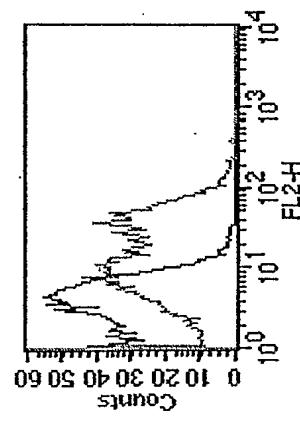
PCT/US05/27239 235/487

19224134 type 4 pilus present in M12

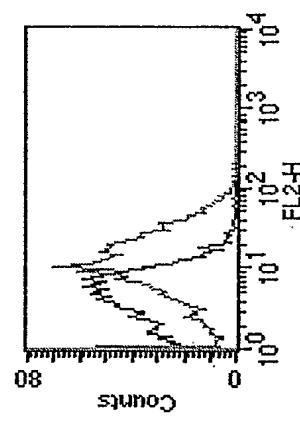
2728 M12



2724 M6



3650 M6



2894 M6

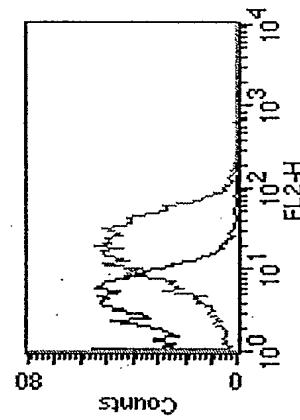


Figure 85

19224135 type 4 pilus present in M12

2728 M12

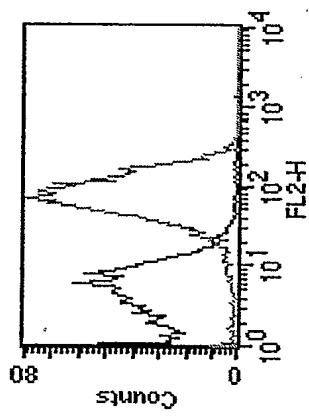


Figure 86

237/487

19224137 type 4 pilus present in M12

2728 M12

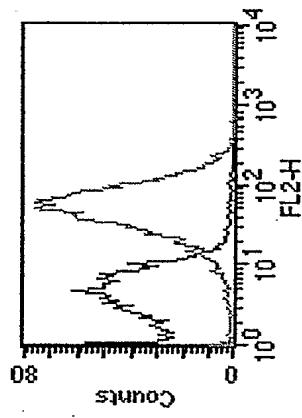


Figure 87

19224141 type 4 pilus present in M12

2728 M12

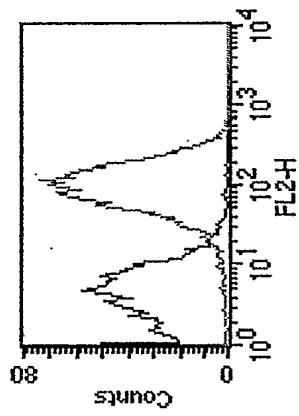


Figure 88

239/487

Figure 89

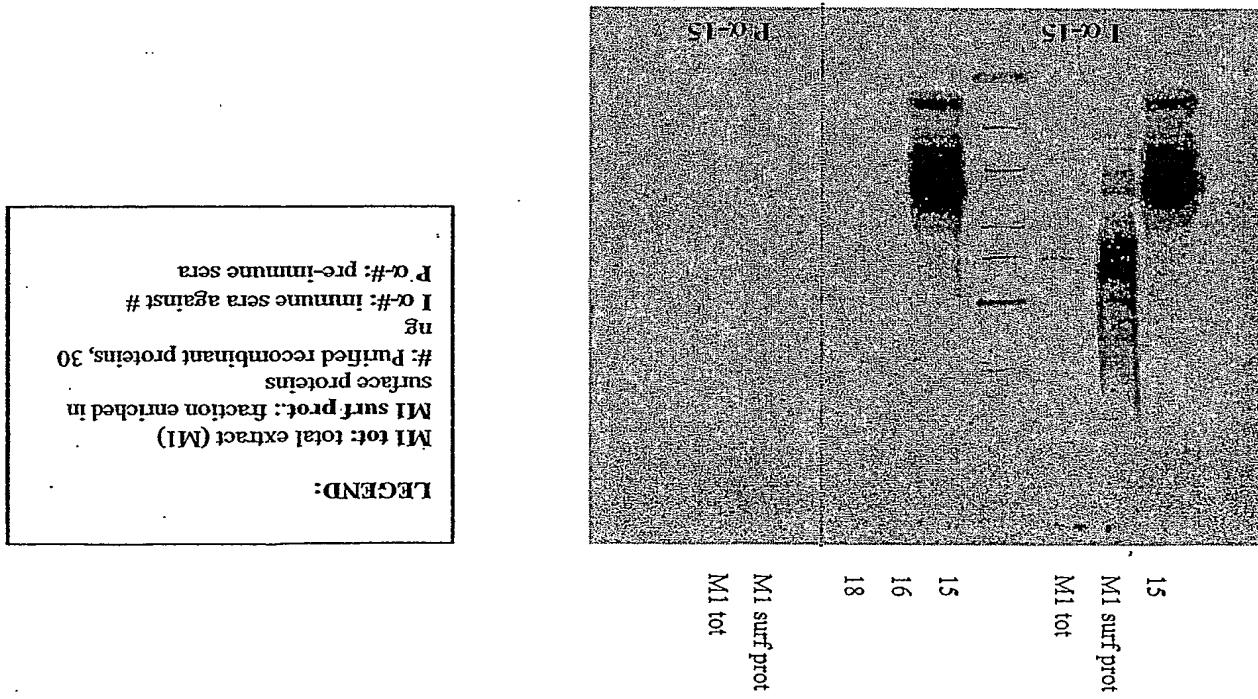
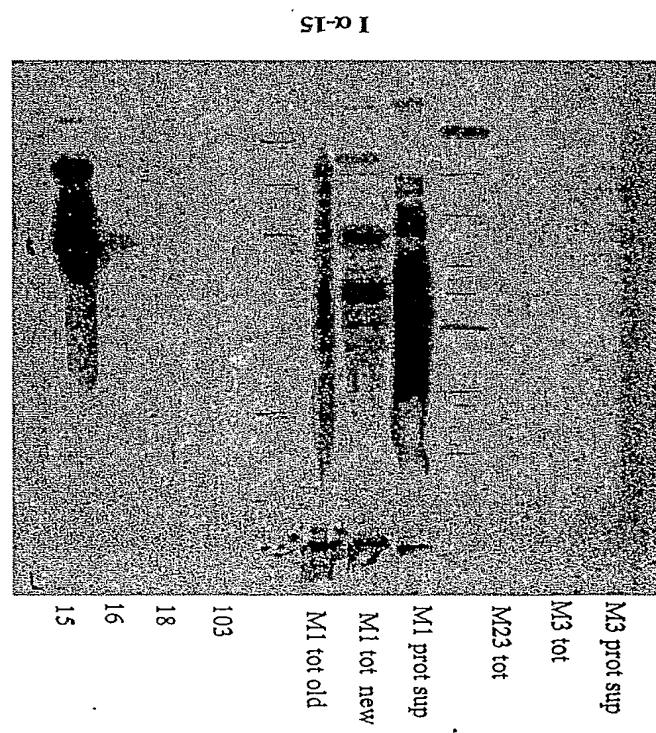
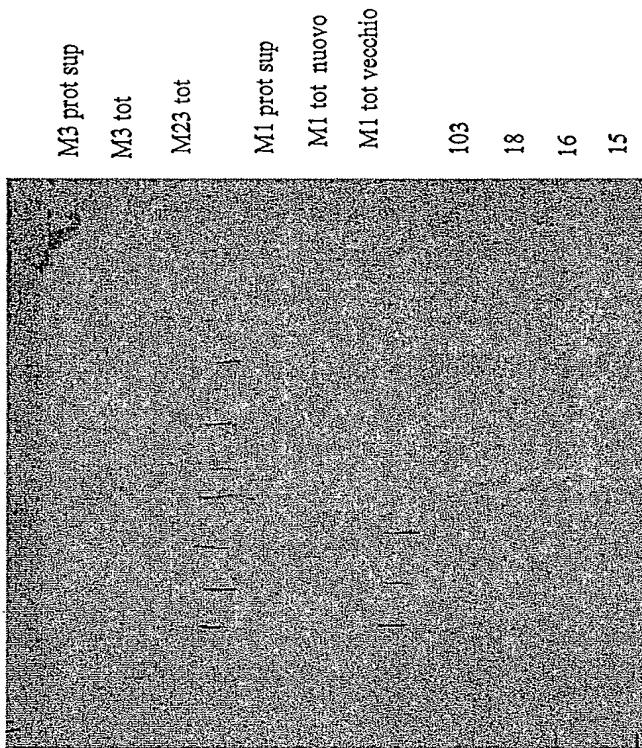


Figure 90

LEGENDA:

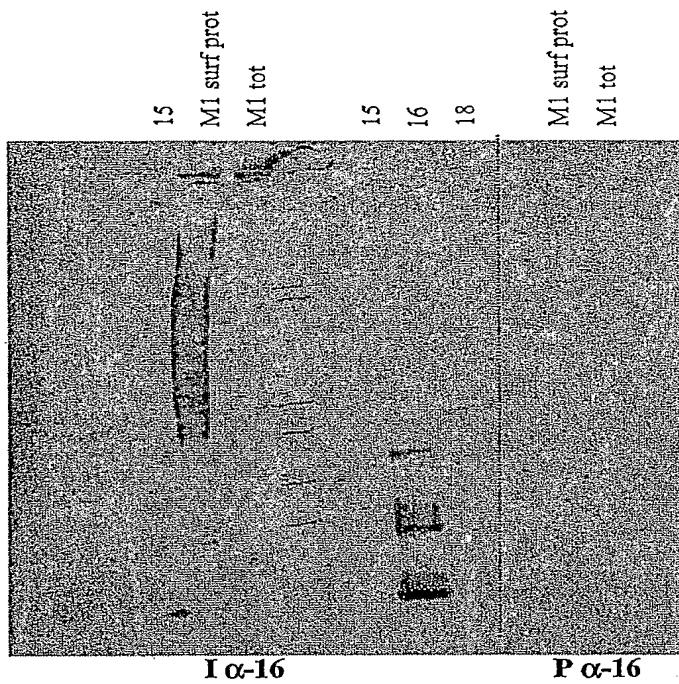
- M1 tot total extract (M1)
- M1 prot Sup: fraction enriched in surface proteins
- #: Purified recombinant proteins, 30 ng
- I α #: immune sera against #
- P α #: pre-immune sera



PCT/US05/27239
241/487 $P \alpha-15$ **LEGENDA:**

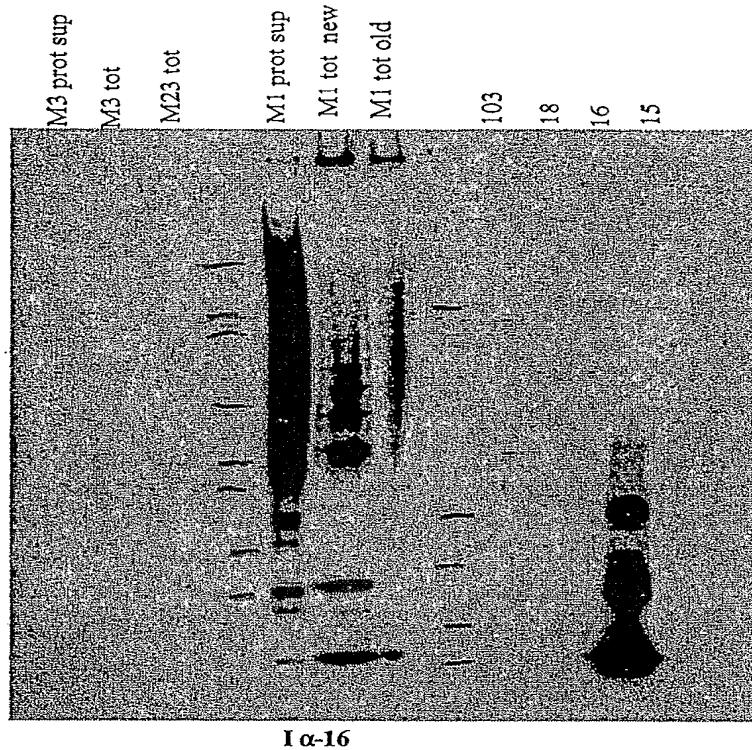
- M1 tot:** total extract (M1)
- M1 prot. Sup.:** fraction enriched in surface proteins
- #: Purified recombinant proteins, 30 ng
- I α -#:** immune sera against #
- P α -#:** pre-immune sera

Figure 91

**LEGEND:**

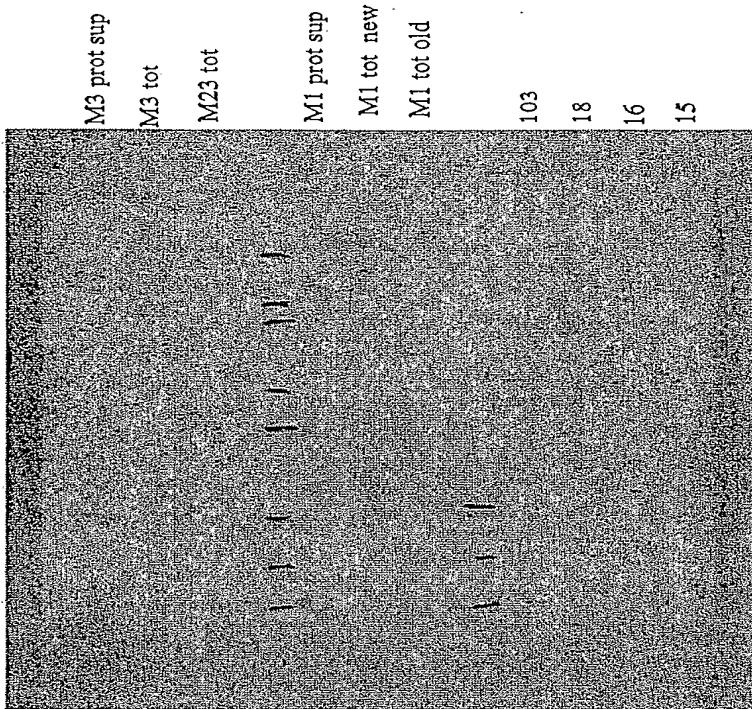
M1 tot: total extract (M1)
M1 surf prot.: fraction enriched in surface proteins
#: Purified recombinant proteins, 30 ng
I α-#: immune sera against #
P α-#: pre-immune sera

Figure 92

**LEGENDA:**

M1 tot: total extract (M1)
M1 prot. Sup.: fraction enriched in surface proteins
#: Purified recombinant proteins, 30 ng
 $I \alpha-\#$: immune sera against #
 $P \alpha-\#$: pre-immune sera

Figure 93

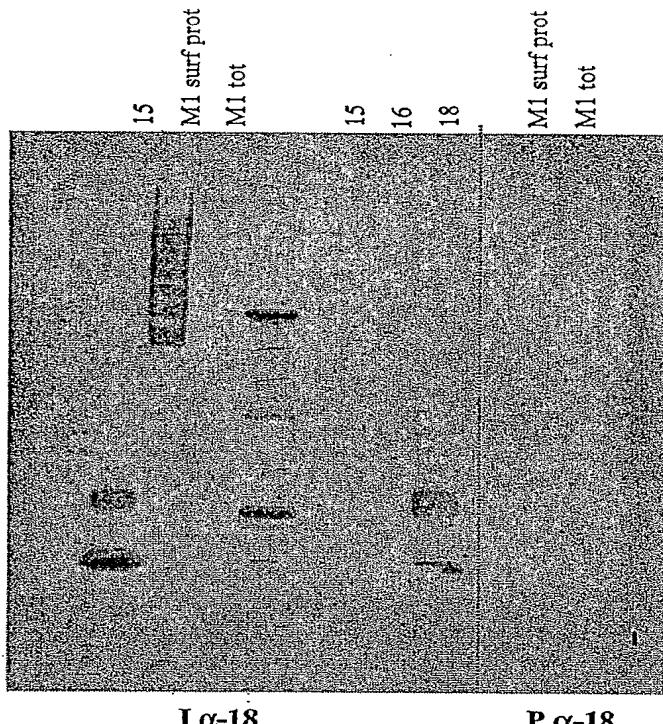


P α-16

LEGENDA:

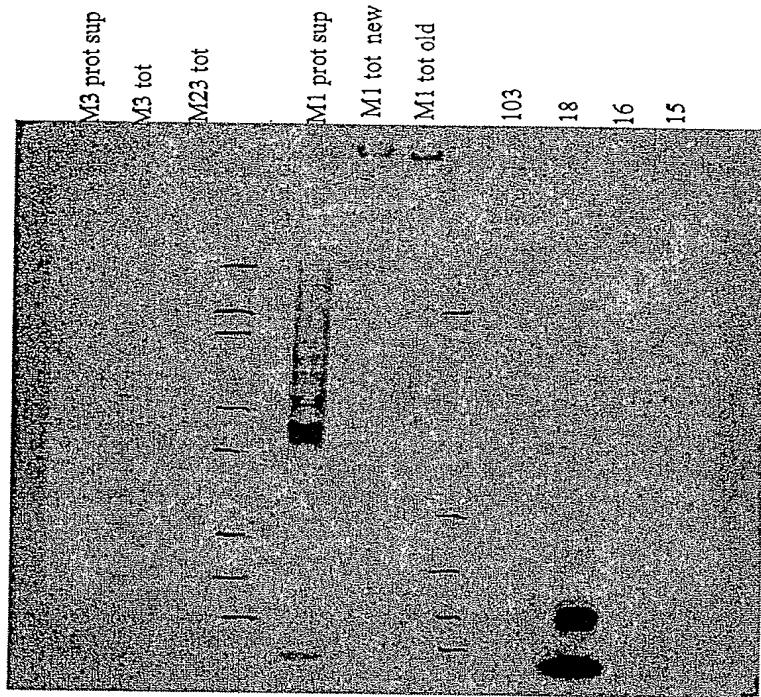
- M1 tot:** total extract (M1)
- M1 prot. Sup.:** fraction enriched in surface proteins
- #: Purified recombinant proteins, 30 ng
- I α-#:** immune sera against #
- P α-#:** pre-immune sera

Figure 94

**LEGEND:**

M1 tot: total extract (M1)
M1 surf prot.: fraction enriched in surface proteins
#: Purified recombinant proteins, 30 ng
 $I \alpha\text{-}#$: immune sera against #
 $P \alpha\text{-}#$: pre-immune sera

Figure 95

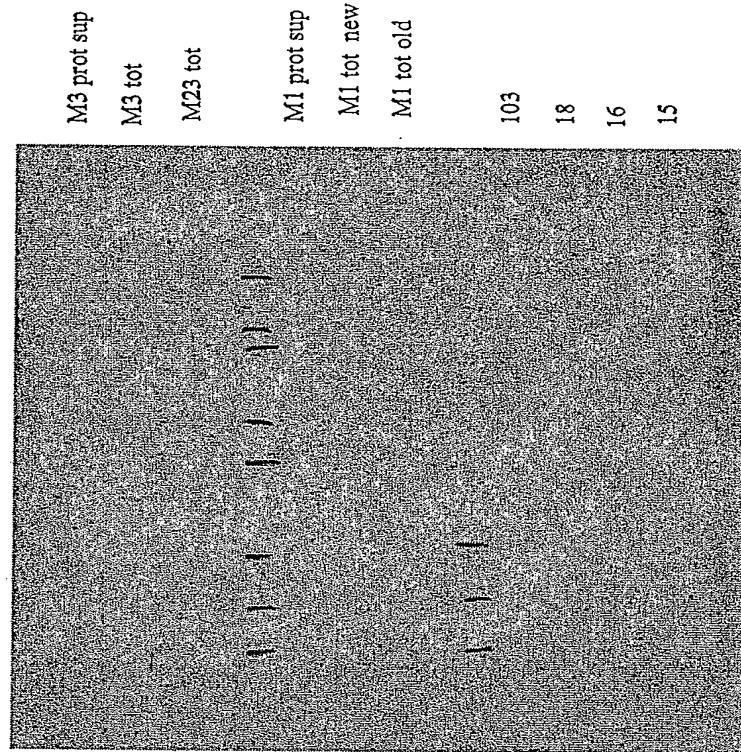


I α -18

LEGENDA:

- M1 tot:** total extract (M1)
- M1 prot. Sup.:** fraction enriched in surface proteins
- #:
#: Purified recombinant proteins, 30 ng
- I α -#:
I α -#: immune sera against #
- P α -#:
P α -#: pre-immune sera

Figure 96

PCT/US05/27239
247/487**LEGENDA:**

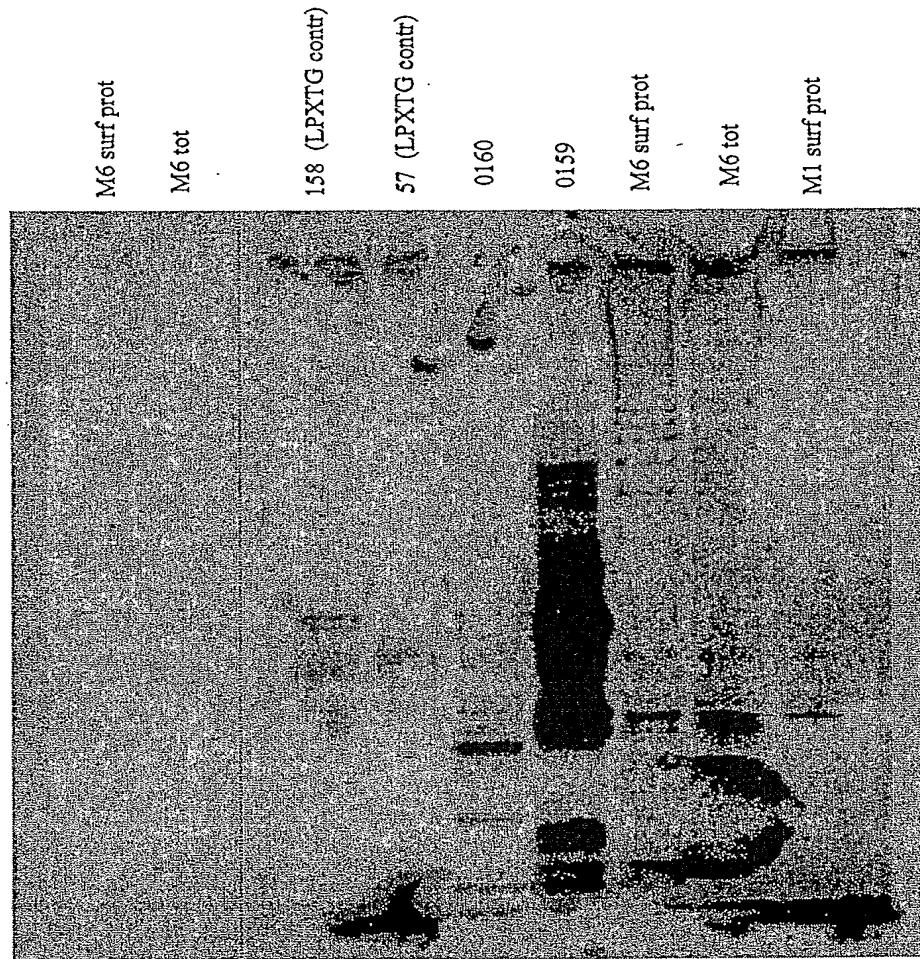
M1 tot: total extract (M1)
M1 prot. Sup.: fraction enriched in surface proteins
#: Purified recombinant proteins, 30 ng
I α -#: immune sera against #
P α -#: pre-immune sera

P α -18

Figure 97

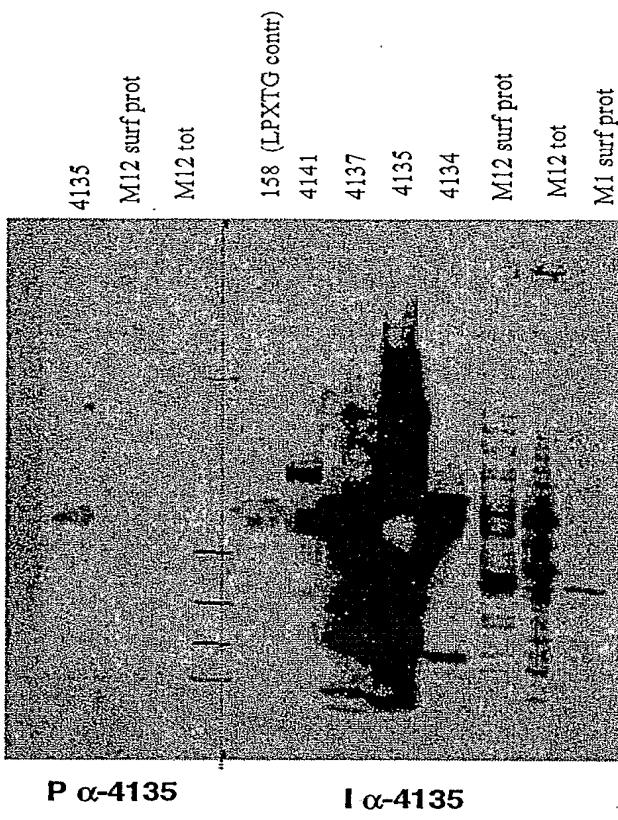
PCT/US2005/027239
248/487

Figure 98

**LEGEND:**

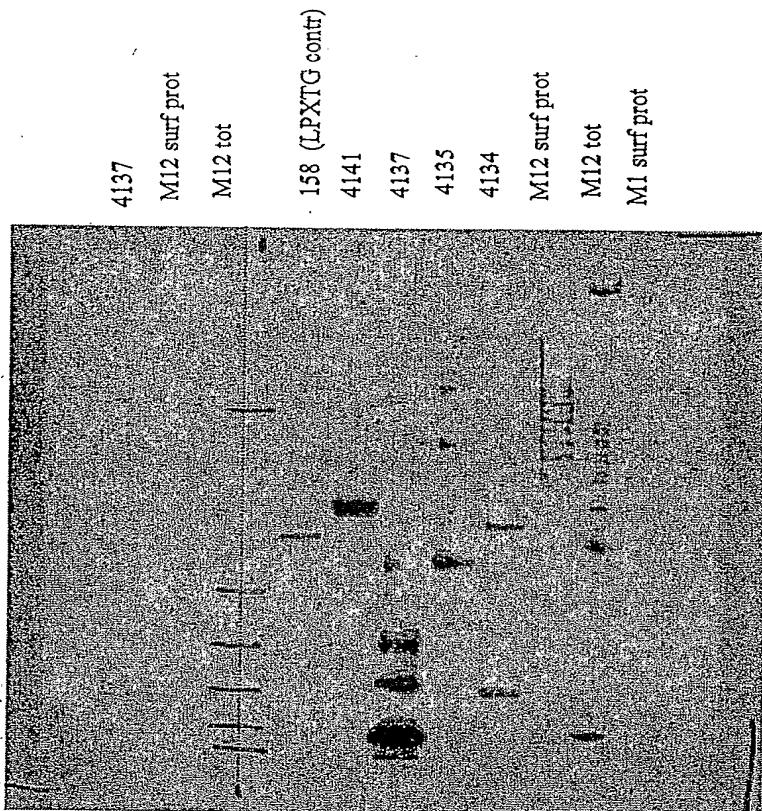
- M6 tot: total extract (M6)
- M6 surf prot.: fraction enriched in surface proteins
- #: Purified recombinant proteins, 30 ng
- I α -#: immune sera against #
- P α -#: pre-immune sera

PCT/X/US05/X27239 249/487

**LEGEND:**

- M12 tot: total extract (M12)
- M12 surf prot.: fraction enriched in surface proteins
- #: Purified recombinant proteins, 30 ng
- I α-#: immune sera against #
- P α-#: pre-immune sera

Figure 99

PCT/US05/27239
250/487**LEGEND:**

M12 tot: total extract (M12)
M12 surf prot.: fraction enriched in surface proteins
#: Purified recombinant proteins, 30 ng
I α-#: immune sera against #
P α-#: pre-immune sera

Figure 100

PCT/US05/E27239
251/487

FIGURE 101

GACAGGCTTCCTTATACGGACCGCTTTCTATAATGGACTCTTCCAAAGTTCTTGCTCTTTTA
 1 +-----+-----+-----+-----+-----+-----+-----+-----+-----+ 60
 CTGTGCAAGGAATATGCTGGCAGAAGATATAGCCTGAGAAGGTCAAGAACGAGAAAT

CCAGGGACTAGOCGTTCAGGTGCACAGATTGCGTGGTGGTTAATGGAACCGAGTCGT
 61 +-----+-----+-----+-----+-----+-----+-----+-----+-----+ 120
 GGTOCTGATGCGCAAGTCCACGTTGCTAACAGCCACCAAAACAATTACCTGGTCAGCA

TCAAGTTGTGACAGAATTACCTTCATCTGGATTCCGTTATGTTGGAGCTAGTGCC
 121 +-----+-----+-----+-----+-----+-----+-----+-----+-----+ 180
 AGTCAACACTGTCTTAAATGGAAAGATAGAACCTAAAGGCCAATACAAACCTCGATCACGG

TAAAAAGATTTCAAATTGTAAGCCGGGAAACTCTTGAGCTTGTGGCAATGTGTTTTG
 181 +-----+-----+-----+-----+-----+-----+-----+-----+-----+ 240
 ATTTCCTAAAGTTAACACTTTCGGCCTCTTGAGAACTCGAACCGTTAACAAAAAC

CTCTTGGTGCATGGGAGTAGCTTGTGGGTCAAGCATGGCTATTGCTCTTGAC
 241 +-----+-----+-----+-----+-----+-----+-----+-----+-----+ 300
 GAGAACCGCGCTACCCCTCATCGAAAACGCCAGTGTGTTACCCGATAAGCGAAACTGG

AGCTATGTGAAAAAACACGACTTCACCCCTTTGGTAAATACCGTATCGTGTGGTAGT
 301 +-----+-----+-----+-----+-----+-----+-----+-----+-----+ 360
 TCGATACACTTTTGTGCTGAAGTGGAAAACCATTATGGCATAGGCACGAACCATCA

GTTTGCTACTTTACAGTTTCGGTTTATTTGTATAAGAAAACCTTGAAGGGTAAC
 361 +-----+-----+-----+-----+-----+-----+-----+-----+-----+ 420
 CAAACGATGAAATGTCAAACAGGCCAATAAACATATTCTTTGGAAACTCCCCATTG

TCTTCAGGGTTTATACCTCTAGAAAATCTCTCAAAACGGCTAGCTTATCTGCAACC
 421 +-----+-----+-----+-----+-----+-----+-----+-----+-----+ 480
 AGAAAGTTCCAAGGAAATATGAGAATCTTTAGAGAAGTTGGCAGTCGAATAGACGTTGG

Figure 101A

TCAAAACAGTGTGTTGAGCAGCCCTGCCATAGTTCTGCCTGAGTCATT
 481 +-----+-----+-----+-----+-----+-----+-----+-----+ 540
 AGTTTGTACAAAACCTCGTGGACGCCATCGAAGGATCAAACGAGAACTAAAGTAA

GAGCTTAAATCAGTCAGGGTAATCCCCAATAGGGACACCTCTTCTTCTGCTT
 541 +-----+-----+-----+-----+-----+-----+-----+-----+ 600
 CTCGAATTGGCTCAGTCCCATTAGGGTTATCCGCCCTGTTGAGAAAGAGCGAA

AATTCTCATAGAGTTGAGGGCTATTGGCTTATCTGACTAGCATCTTGTTTGG
 601 +-----+-----+-----+-----+-----+-----+-----+-----+ 660
 TTAAGAACGATCTAACGTCGCCATAACGAATAGACTGATCGTAGAACCAAAACC

CAAGACCTTTTCGTTGGTAAGAGTTGAAAAGTCCTCTCGTAGGGATTCTAACATGACAA
 661 +-----+-----+-----+-----+-----+-----+-----+-----+ 720
 GTTCGTGAAAAGCAAACCTCTCAACTTTCAAGGACATCGCTAAAAGTTTACTGTT

TTTTCCAGCTTTCTTGTGATGTTGAGTTGAGGGACTTTCTGATAGAGAGTCA
 721 +-----+-----+-----+-----+-----+-----+-----+-----+ 780
 AAAAGGTGAAAAAGAACATCACATCTAACCTCTAACATCTGCTGAAAGACTATCTCAGT

GCTCTTTGATATCTTCCCTGGCACGGAGAACCTCCCGTAGGGTTCTCCTGCGGA
 781 +-----+-----+-----+-----+-----+-----+-----+-----+ 840
 CGAGAAAAAAACTATAAGGGACCGCTGCCCTAGAAGGGCATCCAAAGGGACGGCT

TTGATTTACGGATTGGGATTGGATTGACTGGAGTTGTAATGCCACGGCTTCGAT
 841 +-----+-----+-----+-----+-----+-----+-----+-----+ 900
 AACTAAATGCCTACGCTAACCTAAACTGACCTCTCAACACTTACGGTGGCTGGAAAGCTA

ACAGATCATAGCCTAGTACCAAAACGGCTTACCTCAGGAACCTCAAGTA
 901 +-----+-----+-----+-----+-----+-----+-----+-----+ 960
 TGTCAGTATCGGATCAGATGGTTTGCAGATAATCCAATGGAGTCCTGAGTTCA

253/487

Figure 101B

AATCAGGCCAGGTTAAAACGCCATTGGATGAAAGACGTTCTACTGTCTTCTACTC
961 TTAGTCGGTCAATTGGGTAAACTACTCTGCAAGATGACAGAAAAAGGATGAG 1020

CATGAATTTGAAATATCCATTGGTTGAGAAATCCTCAGCCGTTAGGTAGAATCA
1021 GTACTTTAACCTTTAGGTAAACAAACTCTTTAGGAGTCGGACAAGTCATCTTAGT 1080

CTGTCAAACCATTGGTTTGATAATCACTCGCCATTAGCTAAGAATTGTTGTAAG
1081 GACAGTTGGTACACCAAAACTATTAGTAGGGTAAATCGATTCTAAACACATTC 1140

AAACGCCCTGGGGAGGAGCTTAACTGGGAGTTCTTCCAGATACTCTTTGAATGGGGAG
1141 TTTGGGACGCCCTCGTCAAACTCACCTCAAGAAAGGCTATAGAAAAACTTACTCCGGCTC 1200

CAATTTGACCGCTGACTTGATAACCGAGTTTATTCGTACATCCAATAGGCTTCGT
1201 GTTAAACTGGGACTGAACATGGCTCAAATAAAAGACAGTGTAGGTTTATCCGAAGCA 1260

CAATGCTCATGGGTCAATCAAATCTGTTAGGCCTTAAAATAGCTCGAAATCCGGAGTC
1261 GTTACGAGTACCCAAAGTTAGGTTAGACATATCGCGAATTTCATGGCTTAGGCCCTCAG 1320

CCACAGACTTGTATTCTCATTAATTCCCTGAGATAAAGACAGCCCTGGGGACAAACGTTCAT
1321 GGTGGCTGAACATAAAGAGTATTAAAGGGACTCTATTCTGTCGGACCCCTGTGCAAGTA 1380

AAGCTTCCTGGAACTCATGGCAGAATGGACACCAAAGCTCTTGCCTCATAAACTACAGG
1381 TTCAAGGAACCTTGAGTACGTTACTCTGGTTTACCTGTTGAGAACGGAGTATTGATGTCC 1440

PCT/US05/27239
254/487

Figure 101C

1441 TAGAACGACTCCCCGTCCACCGTTGGCGAGGGTTCCTTCCAAATGACAGGTTTC
ATCTTGCTGAGGGCAGGGTGCACAAACGGCTCCCAGCAAGGTTATTACTGCCCCAAG
CTCTGAGTTAGGATTATCCCTGATTCCACTGCAGCAAAAAGGCATCCATGTCATAAT 1500

1561 GAGACTCAAATCCTAATGGGACTAAAGGTGACGTGTTCTCTAGGTACAGTTATA
GGATGATTTCCTTGACAAATCATTAAACAAAGGAAATCACATGCCCTAGCACCTTTT 1560

CCTACTAAAAGAACTGGTTAGTAATTGGTTCTCTTTAGTTGACGGATGGGGAAA
TATACTCTCGAAAATCTCTCAAAACCACGTCAAGCTTCCATCTGCAACCTCAAAACAGTA 1620

1621 ATATGAGAAGGCTTTAGAGAAAGTTGTCAGTGAGCTGAAGGTAGACGGTGGAGTTTGICAT
TTTTGAGCTGACTCGTCAGTTCTTACAACCTCAAGCAGTGCTTGGAGTTGTCACGAAACTCGTGGAGC 1680

1681 AAAACTCGGACTGAAGCAGTCAGATAAATGTTGGAGTTGTCACGAAACTCGTGGAGC
GGCTAGTTCCCTAGTTGCTTTCGATTTCATTGAGGTGAACTGGCTTATTTCTTTTAT 1740

1741 CCGATCAAAGGATCAAACGAAAGCTAAAGGTAACTCACATGACGAAATAAGAAAATA
TATACCCCTTTCTGAAAAAAGAAAAAGGACTTTTATTTTTCAAAATATAATACA 1800

1801 ATATGGAAAAAGACGTTTCTTCTTCTTCTGAAATAAAAAGTTTATATTATGT
GTTTGAATAATAAGACTGTTTGTAGAAAAGAAAGTGTAAAATAGGAATTTCAC 1860

1861 CAAACTTTATTTATCTGACAAAATCTTTCATTTTCAATTATCCTTAAAGTGA
1920

Figure 101D

1921 TGTTGAAATCGGTTACTTTATGGTATACTTGTCTCATGAATGTAACAGATGACTGTTACT
ACAACTTTAGCCAATGAAATACCATATGAAACAGAGTACTTACATTGTCTACTGACAATGA

1980

1981 AGAAAAGGGACATTAATATGGTTGTTAAGACAGTTGTTGAAGGCACAAAGATAATTG
TCTTTCCTCCTGTAATTACCAACAAATTCTGTCAAACAAACTTCGTGTCATAAAC

2040 M V V K T V V E A Q D I F D - this orf is the homologue of sp0459, a formate acetyltransferase (pf1). It is out of the pilus locus

2041 ACAAGCTTGGAAAGGCTTCAAGGCCTAGATTGAAAGAAAAAGCAAGTGTATCACGCT
TGTTTCGAAACCCCTCCGAAAGTTTCCGCAATTCAACCTTCTCGTTCATAGTGGCA

2100

C K A W E G F K G V D W K E K A S V S R F -

2101 TTGTTACAAGCTTAACACACCTTATGATGGAGACGAAAGCTTCTCTGGACGGACCAACAG
AACATGTTGATGATGTTGAAACTACCTACCTCTGCTTTCGAAGGAACGTCCGGTGTGTC

2160

C V Q A N Y T P Y D G D E S F L A G P T E -

2161 AGGGTTCACTTCACATCAAGAAATTTGAGAAGAAACTAAAGCACACTACGAAACTC
TCGCAAGTGAAGCTGTTCTTTAACATCTTCTTGTGTTGATTTGCTGTGATGGCTCTTGTGAG

2220

C R S L H I K K I V E E T K A H Y E E T R -

2221 GTTTGCCAATGGACACTCGTCCAAACATCTATCGCTGATATCCCTGGCTGGATTATCGACAA
CAAAGGGTTACCTGTGAGCAAGGTGTAGATAAGCGACATATAGGGACCGACCTAAATAGCTGT

2280

C F P M D T R P T S I A D I P A G F I D K -

2281 AAGAAAATGAAAGTTATCTTGGTATCCAAAATGAAACTCTCAAAATTGAAACTTCATG
TTCTTTACCTCAATAGAAACCAATAGTTTACATACCTTGAGAAGTTACTTGAGAAGTACG

2340

Figure 101E

C E N E V I F G I Q N D E L F K L N F M P -
 CAAAAGGTGGTATCCGTATGGCTGAAACTACTTTAAAGAAAAATGGATAACGAAACGCC
 2341 GTTTTCCACCATAGGCATACCGACTTTGATGAAATTTCCTTACCTTATGCTTGGCTGG + 2400

C K G G I R M A E T T L K E N G Y E P D P -
 CAGGTGTTACGAAATCTTCACTTAATAATGTAACAAACAGTTAACGACGGTATTTCGGTG
 2401 GTCGACAAGTGTCTTAGAAGTGAATTACATTTGTTCAATTGGCTGCCATAAAGGCAC + 2460

C A V H E I F T K Y V T T V N D G I F R A -
 CCTTACACTTCAAATATTCTGGTCCGCTCGTCATGCCACACTGTAACTGGTCTTCCAGATG
 2461 GGATGTGAAGTTATAAGCAGGGGAGGAGTACGTGTCGACATGACAGAGGTCTAC + 2520

C Y T S N I R R A R H A H T V T G L P D A -
 CATACTAACGGGAGCATCGGTGTTACGCACCTCTGCTCTTACGGTGCAGACT
 2521 GTATGAGTGGCCTGCAATGAGCCACAAATGCGTCCAGAACGAAATGCCACGTCTGA + 2580

C Y S R G R I I G V Y A R L A L Y G A D Y -
 ACTTGATGCCAAGAAAAGTAAACCACTGGATGCCAATCAAGAAATGCCATGAAACAA
 2581 TGAACTACGGTCTTTCATTGCTGACCTTACGTTAGTTCTTGTACTTCTTGT + 2640

C L M Q E K V N D W N A I K E I D E E T I -
 TCCGGCTTCTGGAAAGTAAACCTTCAATACCAAGCATGGCAACAGTTGCTGCTGG
 2641 AGGGAGAACCACTTCCTTCATTGGAAAGTTATGGTTCTTAACCTTGTCAACAGGGGACC + 2700

C R L R E E V N L Q Y Q A L Q Q V V R L G -
 GTGACCTTACGGGGTGTAGTGTGGCAANACCAGCGATGAAACGTAAGCAATCCAAAT
 2701 + 2760

PCT/US05/27231 257/487

Figure 101E

PC T/US05/E7E3258/487

Figure 101G

-ACAAATTGGCAACACTCCGTACCTGACATAGGCCAACAAACACTTCTTA
3121 -----+-----+-----+-----+-----+-----+-----+-----+
TGTAAACGGTATGTTGAGGCCAGCATACGTTTACTCGGTGTTGTGAGAGAT
3180 -----+-----+-----+-----+-----+-----+-----+-----+

C K L P Y N F R R Y C M H M S H K H S S I -
TICCAAATCGAAGGTGTAACACAATGGCTAAAGACGGATATGGTGAATGGCTGTATCT
3181 -----+-----+-----+-----+-----+-----+-----+-----+-----+
AGGTATGCTTCCACATTGTTTACCGATTCTGCCTTACCACTTACTGACATAGA
3240 -----+-----+-----+-----+-----+-----+-----+-----+-----+

C Q Y E G V T T M A K D G Y G E M S C I S -
CATGCTGTGTGTCACCTTGATCCAGAAAATGAAAGAACACATCAGTACT
3241 -----+-----+-----+-----+-----+-----+-----+-----+-----+
GTACGACACAGAGGTGAAACTAGGTCTTTACTTCTGTGGGTGTAGTCATGA
3300 -----+-----+-----+-----+-----+-----+-----+-----+-----+

C C V S P L D P E N E E Q R H N I Q Y F -
TGGGTCTGTAAACGTTCTTAAAGCCCTCTTACTGGTTGATGGGTACGACG
3301 -----+-----+-----+-----+-----+-----+-----+-----+-----+
AGCCAGGACATTTGCAAGAAATTGGAGAATGACCAAACTTACCAATGCTGC
3360 -----+-----+-----+-----+-----+-----+-----+-----+-----+

C G A R V N V L K A L L T G L N G G Y D D -
ATGTTACAAAGACTACAAAGTTGATATGAACTACCGTGAAGGTCTTGAAAT
3361 -----+-----+-----+-----+-----+-----+-----+-----+-----+
TACAAGTGTCTGATGTTCTATAAAACTATAGCTTGGCTAGCTGCTTCAGAAACTTA
3420 -----+-----+-----+-----+-----+-----+-----+-----+-----+

C V H K D Y K V F D I E P I R D E V L E F -
TTGAATCAGTTAAGCGAACTTGAAAAATCTCTTGACTGGTTGACTGACACTACGTAG
3421 -----+-----+-----+-----+-----+-----+-----+-----+-----+
AACTTAGTCATTGCTTGAACATTAGAGAACTGACCAACTGACTGAAATGCATC
3480 -----+-----+-----+-----+-----+-----+-----+-----+-----+

C E S V K A N F E K S L D W L T D T Y V D -
ATGCCCTTGAACATCATCCACTACATGACTGATAGGTACAACACTACGAAAGCTGTTCAAATGG
3481 -----+-----+-----+-----+-----+-----+-----+-----+-----+
TACGGAAACTTGTAGGTGATGTACTGACTATCCATGTTGACTGCTTCGACAAGTTTAC
3540 -----+-----+-----+-----+-----+-----+-----+-----+-----+

PCT/US05/16505 / 27239
259/487

Figure 101H

c A L N I I H Y M T D R Y N Y E A V Q M A -
 CCTTCCTGCCAACTAAACAACGTGCGAACATGGATTCTGGTATCTGGATTGGCTAACAA
 3541 GGAAGAACGGTTGATTGGTGCACGGTTGACCTTAAGCCATAGAACCTAACGGATTGT
 c F L P T K Q R A N M G F G I C G F A N T -
 CTGGTGTATACTTGTCAAGCTATCAATAAGCTACAGTTAACCAATCCGTGACGAAGATG
 3601 GACAACATATGTAACAGTCGATACTGTTATGGCATGGTCAATTGGTAGGCACITGGTTCTAC
 c V D T L S A I K Y A T V K P I R D E D G -
 GCTACATCTACGATTACGAAACAAATCGGTGACTACCCACGTGGGTGAAGATGACCCAC
 3661 CGATGGTAGATGCTAATGCTTTGGTTAGCCACTGATGGGTGGACCCACTTCTACTGGGTG
 c Y I Y D Y E T I G D Y P R W G E D D P R -
 GTTCAACGAATTGGCAGAATGGTTGATCGAAGCTTACACAACACTGCTTACGTAGCCACA
 3721 CAAGTTGCTTAACCGTCTTACCAACTAGCTTCGAAATGTTGAGCAGATGCTCGGTGT
 c S N E L A E W L I E A Y T T R L R S H K -
 AACTATACAAAGACGCCAGCTACAGTATCACTTTGACAATCACATCTAACGGTTGCTT
 3781 TTGATATGTTCTGGCTCTTGATGTCATAGTGAACACTGTTAGTGTAGATTGCAACGAA
 c L Y K D A E A T V S L L T I T S N V A Y -
 ACTCTAAACAAACTGGTAACCTCACCAGTTACAAAGGGTATAACCTCAACGAAGATGGTT
 3841 TGAGATTGTTGACCATGGTGGTCAAGTGTTCCACATATGGAGTTGCTTCTACCAA
 c S K Q T G N S P V H K G V Y L N E D G S -
 CTGGTGAACTTGGTCTAAACTTGGAAATTCTACCCAGGTGCTAACACAAAGCTAACAAAGCTA
 3901

PCT/US05/27239 260/487

Figure 101I

GACACTGAAAGATTGAACTTAAGAAGAGTGGTCCACGAATGGTAGATTGTTTCGAT
 C V N L S K L E F F S P G A N P S N K A K -
 AAGGTGGTTGGTGCAAAACCTTGAACTCACTTCTAGCCTTGACTTTAGTTATGCCAGCTG
 3961 TTCCACCAACCTGTTGGACTTGAGTGAAGATGGAACTGAATCAAATGTCGAC
 C G G W L Q N L N S L S L D F S Y A A D -
 ACGGTTATCTCATGACTACACAGPATCACCTCGGGCTCTGGTAAGACTCGTGTGATGAAAC
 4021 TGCCATAGAGTAACTGATGTTGTCATAGTTGGAGCGGAGAACCATCTGAGCACTACTTG
 C G I S L T T Q V S P R A L G K T R D E Q -
 AAGTTGATAACTTGGTAACATCCTTGATGGTTACTTGAAAACGGGGACAAACAGTTA
 4081 TTCAACTATTGAACTTGTAGGAACTACCAATGAGGCTTTGCACCTGTGTGCAAT
 C V D N L V T I L D G Y F E N G Q H V N -
 ACTTGAACTTATGGACTTGAACTGTTACGAAAAAAATCATGTCAGGGAAAGACGTTA
 4141 TGAACCTGAAATACTGAACTTGCTACAAATGCTTTTGTAGGAGTGGCTTGTGCAAT
 C L N V M D L N D V Y E K I M S G E D V I -
 TCGTACGTATCTCTGGATACTGTTAACACTAAATACCTCACTCAGAACAAAAACTG
 4201 AGCATGCAATAGAGACCTATGACACATTGTGATTAGGATGAGGTGGCTTGTGCAAT
 C V R I S G Y C V N T K Y L T P E Q K T E -
 AATTGACACAACTGTTCCACGAAGTTCTTCAATGGATGACGCCCTGGATGCAATTGA
 4261 TTAACCTGTTGACAGAAGGTGCTTCAAGAAAGTTACCTACTGGAAACCTAAGTGAACCT
 C L T Q R V F H E V L S M D D A L D A L S -

PCT/US05/22723 filed

261/487

Figure 101J

	GCTAATCAAGTTCTTGGATAATAAAAGGGCTCTTTCGACTGTAGGGTTGAGAAA		
4321	CGATTAGTCAAGAACTTATTATTTCGGAAAACAGTGCACATACCCAACTTCTTT		4380
*			
	AGCTTAAGCTCGAGAAAGGACAATTTCGCTTTCCTTGTGTTTGATGTTCAAGGGATGAAA		
4381	TCGATTGAGCTCTTCTGTAAACAGGAAGAAAAACTACAAGTCTGCTACTTCTT	4440	
e	* A R S L V F K T R E K K I N L A I F -orf1_670 homologue of sp0460, transposase		
	ATCCGTTTTTGAAAGTTCAAGGTTCCGAAACCAAGGCATTGGCCTTGATGTTCTTGG		
4441	TAGGCAAAAACCTCAAAAGTTCAAGGCTTTCGGTTCCGTAACGGAACTACAGAAAC	4500	
e	I R K K F N E F N R F G F A N R K I D K -		
	ATGAGTTTGTAGTGGCCTCAAGTTAGCGTTAGAATAAGGCAATTCAATGGCTTAGTG		
4501	TACTCAAACCAATTCAACGGGAGTTCAATCGCAATTCTTATTCCGTTAACGTACCGCAATCAC	4560	
e	I L K N T A E L K A N S Y P L E I A N T -		
	ATGAGTTTATAGCAAAATAATGCTCAAGTGGTTAAAGGTGGTTGAGATGA		
4561	TACATCAAAAATATCGTTTATTACACGAGTTCAACGAGTTCAACAAATTCCACGCCAACTCTACT	4620	
e	I Y N K Y C I F T S L T T K F T R N L H -		
	GGTAACGGTCTTGTAAATTAGCCCCAAAACCTGGTCAGTATCTCTCTGTAGATGAAT		
4621	CCATTGCAAGAACTTAAATCGGGTTTGACCGACTAGTCATAAGAGAACATCTACTTTA	4680	
e	P L T D Q I L G W F Q D T N K E Q L H F -		
	AGGAGTAGTTGATAACGGTCATAAGTAATCTTAAAGTTCAAGTCTAGGTACTAGAGTAAAGATTTTC		
4681	TCCCTCATCAACTATGGTCCAGTATCAATTAGAATTCAAGTCATGATCTCATTTCTAAAG	4740	

PCT/US605/27239

262/487

Figure 101K

e L I L Q Y I D Y D K E P V L T F I K -
 4741 TTCAGACACTCCCTAGGAGTTAAGGTCTCTGAAGGTTCTAGCATAGAAAGGCTTAAGA
 AAGTCGTGAGGGATCCTCAATTCCAGAGGACTTCAAGATCGTATCTTCCGAATTCT + 4800

e K L C E R P T L T S R F T R A Y F P K L -
 GAGAGTTCCGACTATCCTTTAGGATAAATTTCAGTAAATTAAAGGCTCTGTTATTCC
 4801 CTCTCAAAGGCTGATAGAAAAATTCTATTAAAGGTCAATTAAATTCTCGAGAACATAGG + 4860

e S L K R S D K L I F K W Y Y K L A R Y E -
 AGAGATTATCATCAAATTGCTTCATGATGTTGATTCTAGTCGATTAAGAGCCCTGCAC
 4861 TCTCTAAATTAGTAGTTAACGGAGTACTAACAACTAACGACTAAAGATCAGACTAAATTCTCGGGAG + 4920

e L S K D D F Q K M I N I R T Q N L A R S -
 ATGGTTGGACAATGGGAAACGGATCGAGAACAAATTAGGATTGGAAATTAAATTCTTA
 4921 TACACAACCTGTTACACCTTGTCTTGCTAGCTCTTGTTAAATTCTGAACCCATTAAAGAAT + 4980

e M H Q V I H F R D L V I K A N P F L K K -
 ATGAGGGATAACTTCCAGACATACTAACAGTGACGACTTTAACCTTTCTAGCT
 4981 TACTCTCCCTATATTGAAGGTCTGTATGGTGTCACTGCTGAATTGAAAAAGATGCA + 5040

e I L P I Y S G S M D V T V V K V K R A -
 TCTTTGAGTACTTGAAGAAATGATTGGATGGTTGGCTGTTACAGAACTGGAGAACATAGTCTTAC
 5041 AGAAAGCTCATGAACTTCTTAACTTAAAGCTAACCAACACTGGAGAACATAGTCTTAC + 5100

e E K S Y K F F H N R I T T Q R R N D L I -
 GTCTAGATTTCTAGCTTAACTGCTGAGGAAATGAAACCCATTCCCTCTGGTAG
 5101 + 5160

263/487

Figure 1011

CAGTACTAAAAGAATCACACCTTAAGGACTCGTTACTTTCGGTTAAAGGGAAAGACCATC

e T M I K K T N F D Q A I F A L K G K Q Y -

GAGAATTTCATCCCAGGAGGGATTTCAGGCAAGTGGTGTAAATCCTCTTGAAATGAAT 5161

CTCTTAAGTAGGGTCCCTCTCTAAAGTCCGGTTAACACATTAGGAGAACCTTACTTTA

e S F E D W S L I E P L T T Y D E Q F H F -

TGCCTTGGCTTACCGATAAGCCTAGACCTAGAGCTAGAGCATGGCTTAATTAGAACGGATA 5220

ACGAAACTCGAAATGCTATCTGCATCTCCATCTCCATCTACCGATTAATCTTCGCTAT

e Q K L K R Y V T S T S T S I A L K S A I -

TGTGTAAGGCCCTCTCTGTAGTAGGAGGTGGCAATTTCAGCTACAGTGGACTTCCGAG 5281

ACACATTCTGGAGACAACATCCATCCCAACCGTTAAAGACAGACTGGTAAAGGCTC 5340

e H T L A E R N L L Q A I K Q R V M E S -

ATTTGGCAATTTCAGCTACAGTGGACTTCCGACAGGACTTG 5341

TAAACCGTTAAAAGACTTGCTCTCAACAAAGTCGATGTCAGTGAAGGCTGTCCTGAAC 5400

e I Q C N K Q V L T F E A V T V K R C S K -

CATTGAAATCGTCCTCTTTCAAATGAATGAGGCTAGGAAACCAATCTCGATAAAA 5401

GTAACCTTGTAGAGAAAAAGTTACTTACTTCGATCCCTGGGGTTAGAGCTATTT 5460

e C Q F R R K L H I L S P F G G I E I F -

GGGATTTTAGAAGGCTTTGGAAAGTCGTATTGATTGTTTCCTTACAGTGTTCAT 5461

CCCTAAATCTCGAAAACCTTCAGCATAAACTAAACAAAGGAATGTCACAAATGTA 5520

e P I K S P K Q F D Y K I Q K C H K C -

PCT/US05/27239 264/487

Figure 101M

TTAGGGTGGGTGATAATCAAGTGTAGCGAAGACTTCGATATGGGTATCGTGCTGAATGGCT
 5521 +-----+-----+-----+-----+-----+-----+-----+-----+
 AATCCACCCACTATTAGTTACATCGGTTCTGAAGCTATACCCATAGCACGACTTACCGA
 e K P P H Y D L T A F V E I H T D H Q I A -

 TTATTAAAGGTGATGTTTGTCTTTATTCCGATGAGTAATGGGTATCAGTGTGTT
 5581 +-----+-----+-----+-----+-----+-----+-----+-----+
 AATAAATTCCACCTACAAAACAGAAAATAAGGCTACTCTATTACCCATACTAACACA
 e K N .L T I N K D K I G I L L T T H N I H -

 TCCATAAGATACTTCTAAATGAGTTGTTAGGGCTTTTCATTATAAGCTTATGGACT
 5641 +-----+-----+-----+-----+-----+-----+-----+-----+
 AGGTATTCTPATGAAAGATTACTCAACAAATCCGGAAAAGTAAATTAGAATACCCCTGA
 e E M -

 TTTTTGAACTCAAAAAGCCCTATAATTCTCCACAGGGATTTACCCACTACAGAAATT
 5701 +-----+-----+-----+-----+-----+-----+-----+-----+
 AAAACTATGAGTTTTCGGGATTATTAGAGGTGTCAACCCCTAAATGGGTATGCTTTAAT

 TAGAGCCAGAAAAAACACTTTTGTTCACCTAGCAGAAACTAGAGGCCAAAGTGTTTCT
 5761 +-----+-----+-----+-----+-----+-----+-----+-----+
 ATCTCGGTCTTTTGTGAAACAAAGTGTCTTGTATCTCTCGTCTTCACAAAAAGA

 GTTCAGATTACCCAAAACTGGGAAATAATGGGTATAGAATAGAGATGGCTTAGGAAGCC
 5821 +-----+-----+-----+-----+-----+-----+-----+-----+
 CAAGCTTAAATGGGTTTGCCTTTTAAACCCCTATTCTPATCTTACCGAAATCTCGG
 e M -

 CCTTTTGTAGACAGTACGATGAACTTACAAATAACTGAGCTTTAGCAATC
 5881 +-----+-----+-----+-----+-----+-----+-----+-----+
 GGAAAAACACATCTGCTCATCTGCTTACCTGGAAATAATCGTGTGATAGATGAGA
 e * L L Y H A K K A I -orf2_670 homologue of sp0461, transcriptional
 ATTGGGACCCGTTTGTCAAAGCCTCTTTTGGATATCTACATTGCTGATAGATGAGA
 regulator

265/487

Figure 101N

5941 TAACGGCTGGCAAACAGTTTCGGAGAAAGCCATATAGATGTTAACAGACTATCTACTCT 6000

e M A V R K D F A E K R I D V I T Q Y I L -

CGCTGTGGTACATGCCAATCTAAGGCAATCGTCAAAAAGTGAATGTTCCCTTTGGGA,
6001 GCGACIACCGGAACTGACGTTAGATTCCGGTAGCAGTTTCACTACAAGGAAACCT 6060

e R Q Q S V H L D L A I T L F H H K G K P -

TACTGCCTTAAACGTAAGGCCAGGTTATTCTTCTGTTGTAATAATAATCAATGCGCTCTGTC
6061 ATGACGAAAAATTGCATTCCGTCCATAAGAAACATTATTAGTTACCGAGACAG 6120

e Y Q K K V Y P L Y E K T T I I L P E T -

AAATGCTCCTCTGAAGGGAGGACTAATTAGAATAATTGTATCCTGTAACAGAGGCACT
6121 TTTACGAGGAGACTCCCTCCCTGATTAATCTTAACTAGGACATTGTCTCCGTGTA 6180

e L H E E S P P P S I L I N Y G T V S A V -

TTTGTCAAGTAAATTCCGTAATAATGGACTTTTAAAGTTAACATCTGCTTGATTATT
6181 AACAGTCATTAAAGGCAATTATTACCTGAAATAATTCAAAATTGAGCAACTAAATAAA 6240

e K D T F N R L I I S . K I L N V D A Q N N -

AAATGATAAAATCGGGATAGCAGGTAGTGGAGAAAGATGGTTTCTGTCAAAGTAGAGT
6241 TTTTACTATTAGCCCTATGGCCATCACTCCCTTCTACCAAAAGACAGTTCATCTCA 6300

e L I I F I P I A P L S S F I T E T L Y I -

GAGGAAAGGTAAGCCGATGGCTGGTGGTCAACTCCCTCAATCTCTGCTCAAGTCATCCAC
6301 CTCTTTCCATGCGCTAGGACCAGCTATTGAGGAAGTTAGAAGCAGTCAGTAGGTG 6360

e S F L Y L R H Q D I V G E I K Q E T M , W -

266/487

Figure 1010

TCTTGAAACATTGCTTTCGAAATATGATAACAGTGGCTTGTGCTTCAATCCATAATGT
 6361 AGAACATTGTTAACGAAAGCTTATACTATGTCAACGGAACAGCAGAAAGTTAGGGTATTACA
 e E Q V I A K S I H Y L P K D S E I G Y H -

 TCGTAATAATTATAATAGGAACTAGATTTGTAACCAAACAAAACGTTCTTGTAAAG
 6421 AGCATTAATTATAATTATCCCTTGATCTAAACATTGGTTGGTTGGCAAGAACATTG
 e E Y Y N Y P V L N Q L G F L F T R T L -

 AAAGTCAGTGCTGTTAAAAAGAAAGAGAATTGCAAAATGTCATTCTTAAGATAATTCTTG
 6481 TTTCACTCACGACAATTTCCTCTTCAAGCTTACAGTAAGGATTCTATAAGAAC
 e F T L A T L F S L S N S I D N G L I N K -

 AACTTGGAATAGTAGATGCTTCTTGTATGCTGAAGAATCAGTTGAATAGGTATGAGTC
 6541 TTGAAACCTATCATCTAGAAAGGAAACATAAGCATTCTTAGTCACCTTACACTCAG
 e F K S L L H K G R T H Q L I L Q I T H T -

 TTTTTCTGATTCCATTGCTCTGGAAACGAAAGAATTAGCAGAACATAACCAAA
 6601 AAAAGAACTAAGGTAACAGGAACCCCTTGCCTCTTAATCGTCCTGGTTAGGGTT
 e K K E Q N W K D K S F S S N A S C Y V L -

 AAGATATAATCCAGTTCTCTGAGTAAAGTCATGGGGCATGGCTCTAAGTAAGTT
 6661 TTCTATATTAGGTCAAGAAGGACTCATTTCAACCGTACACCAGATTCATTCAA
 e F I Y D L E E Q T F T M N A H P E L Y T -

 TGGCAATGTTCCATCAAATGGATACTAAAGGGTTTTAACTTTCAAAACTCTTTG
 6721 ACCGGTACAAGGGTAGTTTGGCTATGTATTCTCCAAAAAAATTAAAAGTTGAGAAC

Figure 101P

e Q C H E M L I P Y M F L N K L K E F E K -
 GACTCAGGAACTCAAGTGGAAATTCCCGACGTTCACAGTGCCACTAGTAGATGCTA
 6781 CTGAGTCGCCCTTGAGTTCACCTTTAACGGCTGCACAGGGTTCACTCACGGTGAATCATCGAT

e S E P F E L P F E R R K W T L A V L I S -
 AAATGAAACATACTCGTCAGGGTGTGATTTCATAACAGTTCATGACTGAGTTGAGATTAGAC
 6841 TTTTACTTGTATGAGCAGTCCACACTAAAGATTGTCAAGTACTGACTCAACTCTTAATCTG

e F H V Y E D P T I E L L E H S L Q S N S -
 TGCACAAATCATATGTTGACCCAATCCATACTTCCATCAATTCAAATCTCAATA
 6901 ACGTGTAGTAACTACACACTGGTTAGGTATGAAGGTAGTAAGTTAGTTAGACTTAT

e Q V I M H T V W D M S G D N L D Y I E I -
 CCAAAATGAAACTGGAGGTGCAATTAAAAAACGAAATGCGATATTAGGGACCAACTACT
 6961 GGGTTTACTTTGACCTCCCTACGGTTAAATTGCTTACGCTTAAAGTCCTGGTTGATGA

e G F H F Q L L A I L F R I R Y E P G V V -
 TGATTTTCAAGGTCCAACCTACTGAACGTAGTAACAAGCCACACTTTTGCGTAGG
 7021 ACTAAAAAGTGTCCAGGTTGGATGACTGCTCATGTTCCGGTGAACAGCATGC

e Q N K V L D L G V S R L L G C K Q R V -
 CGGTAGCCCTGTTGGATGGAAATAATACCTTTTGTGTAATTGCTTAAAGCTTGTGATTA
 7081 GCCATGGACAAACGCTACCTTTATGAGAAAACACATTAAAGCAATTGAAACTAAAT

e R Y G T A I S I Y E K Q T F E N F S Q N -
 CCTTGATGAGAAAGGGAGTATTAAAATAGTTGATTTGATGAAACGCTGATGG
 7141

PCT/US05/27239

268/487

Figure 101Q

GGAACATCATCTTTCGCGCTCATAAAAATTTCACCAATAACCAATATTTCGACTACC
 e G Q L L F F R L I K L I T S Q N Y L Q H -
 AAGTAATAATTCTGGTTGATGAGAATGGGTGCGATTAAATTGAACCTGGTGGTATCTAAA
 7201 TTCATATTAAAGCAAACCTTACCTTACCAAGCTTAATTAACTTGAACAACGCATAAGATT
 e F Y N T Q H S H E I L Q V Q Q T D L -
 TTAAATGTCAACTCTCCCTGAATGGTTCTGTAAATTCTGCAAATGCTTAGGAGACTT
 7261 AATTACAGTTGAGAAGGGCTTACAAGAACATTAAAGGACGTTTACGAATCCTCTGAA
 e N F T E E E F T E Q L E Q L I S L L S -
 TTAGATTGTAATGGAGTTAAAGTAGACAGTTCATCTGTTCAATAGACCGAAATATCCAAT
 7321 AATCTAACATTACTTCATCTGTCAAGTAGATCAAGTTATCTGGCTTATAGGTTA
 e K S Q L S T L E D L E I S R I D L -
 AATATATTAAAATGGTAATTTCATCTGTAATTCTTTCAATGGTATTGGTTAGCATA
 7381 TTATATAAATTTCACCTAAATAGACATTAAAGGTTACATAAACAAATCGTAT
 e L I N L I T I K D T I R K E I Y K N L M -
 GTTACCGAATCTTAGTTGCAATAGATAATTTCATATTAAATACAAAAGAAACTAAT
 7441 CAATGGCTTGAATCAACGTTATCTTAAATTAAATTTATGTTCTTGTTGATTA
 e TGTCTGTCAAAAAGGTTGTGGAATTCCGACTTTATGATAAAACAGCATGTAATAAA
 7501 ACAGAACAGTTTCAACACTTAAAGGTGAAATAACTTTTCGTCGTCATTTATTGTT
 7560 GGCATTTAAAGATAGTAATGGATTGGGGAGTTTATGGCTTATTTTTTATTAGA
 7561

Figure 101R

CCGTAAATTCTATCAATTACTATAACCACTTCAAAATAACCGAATTAAAAATAACT

AAATAATTCTTATCAAAATAATGCGTTCTATAAAAATAATGGATAAAAATACTATT
7621 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 7680

TTTATAAAAATAAGTTTATAACGCAAGATAATTCTATACATTATGATAAA
7681 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 7740

ATTCGATACAAAGTGTGTCATTGCCAGGTGAGAAGATAAGCTATAACGCACTTTATACGC
7741 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 7800

TAAACATATGTTAACACAGTAAACGGTCCAACCTCTCTCATCGATAATTGCGAAAAATATGCG
7801 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 7860

TTTTGCTACGGTTGTAGTGACGGATTAACTCACTGGAGATAATTATCAGAAACATAA
7861 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 7920

AAAACGATGCAAACAAATCACITGCCATTAAUTGAGTCACCTTAATTAAATAATGCTTGTATT
7921 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 7980

TGTTCTATGAAATTGCTTAACAGGGAGAACACACATGAAAAGCTAGAAAGATAATTCA
6861 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 7940

ACAGAGATACTTAACTGCAATTAGCTAACTTCACTTCACTGATACTTAGTATCTTCTTAATTGA
b M L N R E T H M K V R K I F Q -orf3_670 homologue of sp0462, LPXTG

GAAGGGCAGTTGGAGACTTCTCTCTATATCTCACTTGACAGCTTCTTCGATAGTTGC
7981 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 8040

CTTCCGCTAACGTTCTGACGACATATAAGCTCAACTTGCGAAAAGAGCTATCAACG
b K A V A G L C C I S Q L T A F S S I V A -

PCT/US05/23239

270/487

Figure 101s

b TTTAGCAGAACCCCTGAAACCACTCCAGGGATAGGAAAAGTAGTGTAAAGGAGAAGG
 8041 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 AAATCGTCCTTGCGGACTTTGGTCAGGTGCTATCCTTTCATCACATAATTCCCTGTCC
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 b L A E T P E T S P A I G K V V I K E T G -
 CGAAGGGAGGCCCTCTAGGAGATGCCGCTTTGAGTTGAAACAAATACGGATGGCAC
 8101 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 GCTTCCTCCTCGCGAAAGATCCTCTACGGCAGAAACTCAACTTTTGTATGCCCTACCGTG
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 b E G G A L L G D A V F E L K N N T D G T -
 AACTGGTTTCGCAAAGGACAGAGGGCGAACACAGGAGAAAGCGATAATTTCACATAAAC
 8161 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 TTGACAAAGGGTTTCCTGTCTCCGCGTTTCACCGCTCTTCGCTATAAAAGTTTGTATTTGG
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 b T V S Q R T E A Q T G E A I F S N I K P -
 TGGGACATACACCTTGACAGAACCTCCAGTTGTTATAAACCCCTCTACTAAACA
 8221 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 ACCCTGTATGTGAACTGTCTTCGGTTGAGGTCAACCAATAATTGGGAGATGATTGT
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 b G T Y T L T E A Q P P V G Y K P S T K Q -
 ATGGACTGTTGAAGTTGAGAAGAAATGGTCGGACGACTGTCTCAAGGTGAAACAGGTAGAAA
 8281 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 TACCTGACAAACTTCACITCTTACCGCTGCTGACAGGTTCCACITGTCCATCTTT
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 b W T V E V E K N G R T T V Q G E Q V E N -
 TCGAGAAGAGGCTCTATCTGACCACTATCCAGTATCCACAAACAGGGACTTATCCAGATGGTCAAAAC
 8341 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 AGCTCTTCCTCCGAGATAGACTGGTCATGGTGTTCCTGTAATAGGTCTACAAAGTTG
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 b R E E A L S D Q Y P Q T G T Y P D V Q T -
 ACCTTATCAGATTAAAGGTAGATGGTGGAAAAAACGGACAGGCACAAAGGGTTGAA
 8401 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 TGGAAATAGTCATAATTAAATTCATCCTACCAAGCCCTGTTGCTGTGTTCCGCAACTT
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

PCT/US05/22323-9

6. *Am. J. Phys. Chem.*, 1886, p. 100; *ibid.*, 1887, p. 100.

271/487

Figure 101T

Figure 101U

TTTTTATAACTTACTGAGTCGAAATAACCTGCAAGCTAGCATGGCTGCAAATGTCCATTG
 b K I L N D S A L W T F D R T T F T A K T -
 TTATAATTATAGCTTTTTAATCTCACATCAGATCCTACTGATAATTCAAACTTAAAGGA
 8881 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 8940
 AATATTAAATATCGAAAAATTAGACTGTTAGGTAGCTAGATAGTTGATAATTCTCT
 b Y N Y S F L N L T S D P T D I Q T I K D -
 TAGGATTCCATCAGATGCAGGAAATTGAACAAGACAATTGATGATAATTGGGGCG
 8941 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 9000
 ATCCCTAACGGTAGCTACGTCTCCCTTAACCTGTTCTGTTAACTACATACTAGTTAGGCCGCG
 b R I P S D A E L N K D X L M Y Q F G A -
 GACMTTTACCCAGAAGGCTTGTGACCGCTGATGATACTCTGACAAAGCAGGAAAGACC
 9001 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 9060
 CTGAAAATGGGTCTTCCGAAACTACTGGCGACTACTATAGAACTGTTCTGTCGGTTCTGG
 b T F T Q K A L M T A D D I L T K Q A R P -
 AACAGTAAAAAGGTTATTTTCCACATTACAGATGGTGTCCGACTATGTCATAATCCAAAT
 9061 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 9120
 TTTGTCAATTTTCAATAAAAGGTGTAATGTCACCAAGGCTGATACTAGGTAA
 b N S K K V I F H I T D G V P T M S Y P I -
 TAATTTAAATACAGGAACGACGCAATCGTACAGAACTCAGCTGATAATTAAAGC
 9121 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 9180
 ATTAAAATTTATGTCCTGCTGCGTTAGCATGTCCTGACTGATAATTAAATTTCG
 b N F K Y T G T T Q S Y R T Q L N N F K A -
 AAAACTCCAATATAGTACGGGGATAATTACTGGGGACTTGTACATGGTACGGAGATGG
 9181 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 9240
 TTTTGAGGTTTATCATGCCCTATAATGACCTCCTGAAACATGGTACAGTGTCTTAC
 b K T P N S S G I L L E D F V T W S A D G -

PCT/US05/27239

273/487

Figure 101V

b TGAACATAAGATTGTTGGAGATGGAAAGTTATCAGATGTTACGAAGAACCTGT
9241 ACTTGTATTCTAACAGAACCTTCAACCCTTCAATTAGTCCTACAATGCTTCTTGGACA + 9300

b E H K I V R G D G E S Y Q M F T K K P V -
9301 AACAGACCAAATACGGAGTTCATCAAAACTTCAATCACCTCCATGAGCAGAGCTAA + 9360
TTGTCTGGTTATGCCTCAAGTAGTTATGAAAGTTAGTGGAGGTACCTCGTCTCGATT

b T D Q Y G V H Q I L S I T S M E Q R A K -
9361 ATTAGTTTCAGGGGATATAGGTTCTATGGAACTGACTGTATTATATTGGGTGATAG + 9420
TAATCAAGTCGCCCTATCCAGATAACCTTGACTGACATAATAAACCGCACTATC

b L V S A G Y R F Y G T D L Y L Y W R D S -
9421 TATTCTAGCCTATCCATTAACTCTAGTACCGATTGGATAACCAACCATGGTGAACCTAC + 9480
ATAAGATCGGATAGGTAAATTGAGATCATGGCTAACCTAATGGTGGTACCTGGGATG

b I L A Y P F N S S T D W I T N H G D P T -
9481 GACTTGGTATTATAACGAAATAATGGCTCAGGATGGCTATGATGTCCTACTGTGGGT + 9540
CTGAACCATTAATATGGCTTTTACCGAGTCTACCGATACTACAGAAGTGAACCCCCA

b T W Y Y N G N M A Q D G Y D V F T V G V -
9541 TGGTGAACGGGATCCTGGATGAGCAACGGCTACTAGATTATGGAGAGCAT + 9600
ACCACATTGGCCCTAGGACCATGCCCTACTCTGGTGGCGATGATCTAAATACTCGTA

b G V N G D P G T D E A T A T R F M Q S I -
9601 CTCTAGTTCTCTGACAACCTACACTAACGTAACGAGATCCTCAGATTITACAAGAATT + 9660
GAGATCAAGGGGACTGTTGATGCTAGGTAGAGTCAAGTCTAAATGTCCTAA

Figure 101W

b S S S P D N Y T N V A D P S Q I L Q E L -
 GAATCGCTACTTCTATACTATGCTCAATGAGAAGAAATTCTATGAAAATGGTACCGATTAC
 CTTAGCGATGAAGATAATGATAGCAGTTACTCTTCTATGCTTTACCATGGCTAATG 9720

b N R Y F Y T I V N E K K S I E N G T I T -
 AGACCCGATGGGTGAACTAATTGATTTCCAATTGGGAGCAGATGGAAGGTTTGATGCCAGC
 TCTGGGCTACCCACTTGATTAACCTAAAGTTAACCTCTACCTTCACCTCAAACTAGGTGCG 9780

b D P M G E L I D F Q L G A D G R F D P A -
 GGATTACACTTTAACTGCACAAAGATGGTAGTTGTTGGGAAATAATGTCCTACTGGGGG
 CCTTAATGTTGAAATTGACGTTTGCTACCATCAAACCACTTATTACAGGGATGACCCCC 9840

b D Y T L T A N D G S S L V N N V P T G G -
 ACCACAAATGATGGCTGGCTTGCCTAAAAATGCAAAGGGTTCTATGATAACGACTGAGAA
 TGGTGTCTTACTACACCGAACGATTTTACAGTTTACAAGATACTATGCTGACTCTT 9900

b P Q N D G G L I K N A K V F Y D T T E K -
 AAGGATTCGTTGTAACAGGGTTGGCTACCTTGGAACGGGTGGAAAAAGTTACATTGACTTATAA
 TTCCCTAAGCACATTGTCACAAACATGGAACCTTGCCCACCTTTCAATGTAACCTGAAATT 9960

b R I R V T G L Y L G T G E K V T L T Y N -
 TGTTGCTTGAAATGACCAATTGTAAGCAAATAATTCTATGACACGAAATGGTCGAAACAC
 ACAAGGCAACTTACTGGTTAACATTGCTTAAAGATACTGTTACCGCTTACCGCTTGTTC 10020

b V R L N D Q F V S N K F Y D T N G R T T -
 CCTACACCCCTAAGGAAGTAGAAAAGAACACAGTGGCGGACTTCCGATTCCTAAGATTCCG 10080

Figure 101X

GGATGTGGGATTCCCTCATCTTCTGTCACGGCTGAGGGCTAAGGTTCTAAAGC

L H P K E V E K N T V R D F P I P K I R -

10081 TGATGAGAAAGTATCCAGAAATCACAAATTCCAAAAGAGAAAAACTGGTGAATTGAA
ACTACATGCTTCAAGGTCTTAGGTCTTAGGTAAAGGTTCCTCTTTTGAAACCCTTAAC

D V R K Y P E I T I P K E K K L G E I E -

10141 GTTTATAAGATCAAATAAGAAATGATAAAAACCCTGGAGAGATGGGGTCTTGTGTTCA
CAATYATTCTAGTTATTCTRACTATTTTGGTGTACTCTCTACGCCAGAAATCGAAGT

F I K I N K N D K . K P L R D A V F S L Q -

10201 AAAACAACTCGGATTATCCAGATATTATGGCTATTGATCATAATGGCACTTATCA
TTTGTGTAGGCCCTATAGGTCTATAACCTGATACACTAGTTTACCGTGAATAGT

K Q H P D Y P D I Y G A I D Q N G T Y Q -

10261 AAATGTGAGAACACGGTGAAGATGGTAAGTGTGACCTTAAAATCTGTCAGATGGAAATA
TTTACACTCTGCCCCACTTCACTTCAACTGGAAATTTTAGACAGTCTACCTTTAT

N V R T G E D G K L T F K N L S D G K Y -

10321 TCGATTATTGAAATTGACCAAGCTGGTATAACCCGTTCAAATAAGCTATCGT
ACCTATAACCTTCAAGCTGTGACCTTCACTTGGCAAGTTTATCGGATAGCA

R L F E N S E P A G Y K P V Q N K P I V -

10381 TGCCTTCAAAATAGTAAATGGAGAGTCAGAGATGACTCAATCGTTCACAGATAT
ACGGAGTTTATCAATTACCTTCACTGTTAGCAAGGTGTCTATA

A F Q I V N G E V R D V T S I V P Q D I -

PCT/US05/027239 276/487

Figure 101Y

b

10441	ACCAAGGGTTACGAGTTACGAAATGATAAGCACTATATCACAAATTGAGCCAAATTCCCTCC	
	TGGTCGCCAATGCTCAAATGGCTTACTATCGTGTATAATAGTGTGTACTCGGTAAAGGAGG	+ 10500
b	P A G Y E F T N D K H Y I T N E P I P P -	
10501	AAAAAGAGAATACTCCTCGAACCTGGTGTATCGGAATGGTGCCTATCTGATAGGTG	+ 10560
	TTTTTCTCTTATAGGAGCTTGACCATAGCCTTACAAACGGTAAGATACTATCCAAC	
b	K R E Y P R T G G I G M L P F Y L I G C -	
10561	CATGATGATGGGGAGGAGTTCTATTATACACACGGAAACATCCGTAAGGTGTAGCAATGAG	+ 10620
	GTTACTACTACCCTCCTCAAGATAATAATGTTGTGCCTTTGTAGGCCATTACATCGTTACTC	
b	M M G G V L L Y T R K H P *	
10621	AAATGATAATACTCGATACTCTGAGCGATACTTTAAGGAAGTAGCACTCAAGAAGAGATT	+ 10680
	TTTACTATTAGCTATGAGACTCGCTATGAAAATTCTTCATCGTAGTTCTCTCTAA	
10681	AAGTTTACTTGGTGAACAGTTTCTTGCCAAAGTAACCACCATGGAAAGGGGAGATG	+ 10740
	TTCAAATGAAACCCTTTGTCAAAAGAAGGGTTCAATTGGGTAACTTTCCCTCTAC	
10741	TTTTGAAAAACTTGCACAGAAAAGGATATTATTGTCATGTAAATTCAATTACATTGC	+ 10800
	AAAAGCTTTGAACTGTCCTTCTCTAAATAACAGTACACATTAAGTAAGTAAACG	
10801	TCACAGTTGATTAAAGAGATAATGAAATAAGGAGAAATCATGAAATCAATCAACAAATT	+ 10860
	AGTGTCAACTAAATTCTCTATACTTATTCCCTTTAGTACTTTAGTTAGTTAGTTAGTTAAAA	

c

M K S I N K F L - orf4_670, homologue of sp0463, LPXTRG
TAACATATGCTTGTGCTTATTACTGACAGCGAGTAGCCCTGTTAGCTGCAACAGTTT

277/487

Figure 101Z

10861 ATTGTACGAAAGGACCGAATAATGACTGTGGTCATGGACAAAGTGCAGTTGCAAA
 C T M L A A . L L T A S S L F S A A T V F -
 TTGGGGGACAATGTTAGTACAGCACCAGATGCTGTACTAAAACCTTAACATCCATA
 10921 AACGCCGCGCTGTTACAATCATGTCGGCTACGACAATGATTGAAATTGTAGGTAT
 C A A D N V S T A P D A V T K T L T I H K -
 AGTTTACUTGCTTCAGAAGATGATTAAAAGACUTGGGATACAAACGGTCCCTAAACGGATATG
 10981 TCAATGACGGAGAGCTCTACTAAATTTCGAACCCATGTTGCAGGATTCTCTATAC
 C L L E D D I K T W D T N G P K G Y D -
 ATGGAACCTCAATCTAGTTAACAGATTAACTGGAGTTGACTGAGGAATTCCAATG
 11041 TACCTTGAGTTAGATCAAATTCTAAATTGACCTCAACATCGACTCCCTTAAGGTTAC
 C G T Q S S L K D L T G V V A E E I P N V -
 TATACTTTGAAATTACAAAAAGTATAATTGACTGATGGTAAGAAAAAAATCTTAAG
 11101 ATATGAAACTTAATGTTTCATATTAACCTGACTACCAATTCCCTTCTCTTAAGATTTC
 C Y F E L Q K Y N L T D G K E K E N L K D -
 ATGATAGTAAATGGACAAACAGTTCACTGGGGTTGACAACTAAAGATGGACTTAATG
 11161 TACTATCATTTACCTGTTGTCAGTACCCAAACTGTTGATTTCACCTGAAATTAAC
 C D S K W T T V H G G L T T K D G L K I E -
 AACCCAGTACTCTTAAGGTGTGTATCGTATTGCTGAGGATAGAACAAAGACTACCTATG
 11221 TTTGGTCATGAGAATTCCACACATAGCATAAAGCACTCCATCTGTTCTGATGGATAC
 C T S T L K G V Y R I R E D R T K T Y V -

PCT/US605/027239

278/487

Figure 101AA

PCT/US05/27239 279/487

Figure 101AB

C G F N L K L T E A G L A K I N G K D A D -
 ACCAAAAATCCAAATTACTTACTCAGCTTACTTTGAACTCACTTGGTGTGAGACATT
 11701 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 11760

C Q K I Q I T Y S A T L N S L A V A D I P -
 CTGAAAGTAAGGAAATTACATTAATTCAGAAATCATCAAGATCATGGAAATACTCCAA
 11761 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 11820

C E S N D I T Y H Y G N H Q D H G N T P K -
 AACCAACTAAACCDAAATAATGGTCAAATTACAGTAACDAAAGACATGGACAGTCAACCTG
 11821 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 11880

C P T K P N N G Q I T V T K T W D S Q P A -
 CTCCTGGGGTAAAGCGACTTCAACTTGTAAATGCCAAGACTGGTGAGAAAGTCG
 11881 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 11940

C P E G V K A T V Q L V N A K T G E K V G -
 GTGCCTCCCTGTAGAACTTCAGAAATAATGGACATATACTTGGAGTGTAGATAATT
 11941 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 12000

C A P V E L S E N N W T Y T W S G L D N S -
 CTATTGAAATACAAAGTTGAAGGAAATAATGGATACTCAGCTGAATAACAGTAGAGA
 12001 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 12060

C T E V K V E E E V N G Y S A E V T V E S -
 GCAGGGAAAGTGGGGTAAANAAACTGGAAAGATAATACCCTCACATCCTG
 12061 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 12120

PCT/LIS05/27239

280/487

Figure 101AC

CGTTTCCTTCAACCCCCATTTCGACCTTCTATTGGGTCAAGGGTTAGTTAGGAC
 C K G K L G V K N W K D N N P A P I N P E -
 AAGAACACGTTAAACATACGGTAAAAGTTGTCAAAGTAGACCCAAAAGATACTC
 12121 TTCTTGTTGCACATTTGTATGCCATTTCACAGTTCATCTGGTTTCATCTATGAG
 C E P R V K T Y G K K F V K V D Q K D T R -
 GTCTAGAAAATGCCAGTTGTTGTTAAAAAGCAGATAGCAATAAATATTGCCCTTA
 12181 CAGATCTTTACGGTCAAGCAACAATTGTTCTATCGTTATTATAACGGAAAT
 C L E N A Q F V V K A D S N K Y I A F K -
 AGTCAACTGACACAAAGCTGCAGATGAAAAAGCAGCACAACGTGCAAAACAAAAATTGG
 12241 TCAGTTGACGTGTTGTTGACGGCTACTTTCTGTGCGTTGACGTTTGTTTAACC
 C S T A Q Q A A D E K A A T A K Q K L D -
 ATGCAGGGTAGCAGCTTACACAAATGCTGCAGATAAGCAAGGCCGCTCAAGCTCTAGTAG
 12301 TACGTGCCATCGTGAATGTTACGACGTCTATTGTTGAGTCGAGATCATC
 C A A V A A Y T N A A D K Q A A Q A L V D -
 ATCAAGCACAGCAAGAATAACAAATGCTACAAAGAAGCCAAATTGGTTATGGTTGAAG
 12361 TAGTTGTTGTTACATCGAATTGTTCACTGGTTAAACCAAAATACAAACTTC
 C Q A Q Q E Y N V A Y K E A K F G Y V E V -
 TAGCTGGAAAAGATGAAGCAATGGTTCTTACTCTTAATACGGATGGTCAAATTCCAATTT
 12421 ATCGACCTTTCTACTTCGTTACCAAGAATGAAGATTGCTTACCAAGTTAAGGGTTAAA
 C A G K D E A M V L T S N T D G Q F Q I S -

PCT/US05/2723281/487

Figure 101AD

PCT/US05/27239 82/487

Figure 101AE

a V M A L C F S L V W G H A V Q A Q E D -
 CACACGGTGGTCCTGCAATTGGAGAACTATCAGGAGGTCTCCACCAATTCAGTTAACGGTAAACGGCA
 129001 GTGTGCAACCAGAACGGTTAACCTCTTGATAATGTCATGCCATTACGGTTAACGGTAAACGGTAAACGGCA 12960

a H T L V L Q L E N Y Q E V V S Q L P S R -
 GATGGTCATCGGTGGCAAGTATGGATGGATGATTCTATGATGATGGGTG
 12961 CTACCCAGTAGGCCAACGGTTCAACCTACTAAGCATAAAGGATAACTACAGCCCCAC 13020

a D G H R L Q V W K L D D S Y S Y D D R V -
 CAAATTGTAAGAGACATTCGATTGGTGGGATGAGAATAAACTTTCTTCATAAAAGACT
 13021 GTTTAACATTCTCTGAACGTAAGCACCCTACTCTTATTGAAAGAAAGTTTCTGA 13080

a Q I V R D L H S W D E N K L S S F K K T -
 TCGTTTGAGATGACCTTCCTTGAGAATCAGGATTGAAAGTATCCTCATATTCCAAATGGCTTT
 13081 AGCAAACTCTAATCTGAAGGAACCTTACTCTTAACCTCATAGAGTTAAAGGTAAACAGAA 13140

a S F E M T F L E N Q I E V S H I P N G L -
 TACTATGTTGCCTCTTATCCAGACGGATGGGATGGGTTCTTATCCAGGTGAATTTCCTTT
 13141 ATGATACAAGGGAGATAATAGGTCTGCCTAACGCCAACGAAATAGGTCAACTTAAGAAAAA 13200

a Y Y V R S I I Q T D A V S Y P A E F L F -
 GAAATGACAGATCAAACGGTAGAGGCCCTTGGTCATTGAGCGAAAAAAACAGATACATG
 13201 CTTTACTGTCTAGTTGCCATCTGGAAACACAGTAACATCGCTTTTTGGCTATGGTAC 13260

a E M T D Q T V E P L V I V A K K T D T M -
 ACAACAAAGGGTAGGGCTGATAAGGTGGATCAAGGCAACATCGCTTGGGGTGCAGGC
 13261 13320

PCT/US05/27239 283/487

Figure 101AF

a T T K V K L I K V D Q D H N R L E G V G -
 TTTAAATTGGTATTCAGTAGGAGATGGTCTGAAAAGAGGTTCCTGATTGGAGA
 13381 +-----+-----+-----+-----+-----+-----+-----+-----+ 13380
 AAATTAAACCATAGTCATCGTTCTACCAZAGACTTTCTCCAAAGGAACTAACCTCTT

a F K L V S V A R D G S E K E V P L I G E -
 TACCGTTACAGTTCTTCTGGTCAAGTAGGAGAACTCTATACTGATAAAAATGGAG
 13381 +-----+-----+-----+-----+-----+-----+-----+-----+ 13440
 ATGGCAATATGTCAAAGAACGAGTTCAATCCCCTCTTGAGAGATATGACTTTTACCTCTC

a Y R Y S S S G Q V G R T L Y T D K N G E -
 ATTTTTGTGACAAMCTTCTCTTGGAAACTATCGTTCAAGGAGGTGGGCCACTGGCA
 13441 +-----+-----+-----+-----+-----+-----+-----+-----+ 13500
 TAAAAACACTGTITAGAAGGAAACCCTTGATAGCAAAGTTCTCCACCTCGGTGACCGT

a I F V T N L P L G N Y R F K E V E P L A -
 GGCTATGCTGTACGAGCTGGATAACGGATGTCAGCTGGTAGATCATCGCTGGTGACG
 13501 +-----+-----+-----+-----+-----+-----+-----+-----+ 13560
 CCGATAAGACAATGCTGGCAGCCTATGCTCAAGGTGACCATCTAGTAGTCGACCACTGC

a G Y A V T T L D T D V Q L V D H Q L V T -
 ATTACGGTTGTCAATTAGAAATTACCAACGTGGCAATGTCAGTTTATGAAAGGTGGATGGT
 13561 +-----+-----+-----+-----+-----+-----+-----+-----+ 13620
 TAATGCGAACAGTTAGTCCTTATGGCACCGTTACAACCTGAAATACTTCCACACTACCA

a I T V V N Q K L P R G N V D F M K V D G -
 CGGACCAATACCTCTCTTCAGGGCAATGTCAAAGTCATGAAAGAAGAAAGGGACAC
 13621 +-----+-----+-----+-----+-----+-----+-----+-----+ 13680
 GCCTGGTTATGGAGGAAGTTCCCGTTACAAAGTTCAGTACTTTCTGGCTGTG

a R T N T S L Q G A M F K V M K E S G H -

Figure 101AG

13681 - T A T A C T C C T G T T C T T C A A A T G G T T A G G N G T A G T T G T A C A T C A G G G A A A A G T G G T C G T
 AT A T G G A C A A G T T T A C C A T T C T T C A A C A T T G T A G T G C C T T T C T A C C A G C A + 13740

 13741 - T T C C G A T G G A A G G T C T A G G A C T A C T A C T A T T T A T G G G A C T C C A A G T C C A A C T
 A A G G C T C A C C T T C C A G T C A T A C C C T G T A T G T A T A A T C C C T G A G G T T C G A G G T T G A + 13800

 a - F R V E G L E Y G T Y Y L W E L Q A P T -

 13801 - G G T T A T G T T C A A T T A C A T C G C C T G T T C C T T A C A A T C G G G A A A A G G A T A C T C G T A A G G A A + 13860
 C C A T A C A G T T A T G T A T G G G A C A M A G G M A T G T A G C C C T T T C T A T G A G G C A T T C C T T

 a - G Y V Q L T S P V S F T I G K D T R K E -

 13861 - C T C G T T A C A G T G G T T A A A A T A C A A T C A A G C G A C C A C G G A T T G A T G T C C C A G A T A C A G G G G A A + 13920
 G A C C A T T G T C C A A T T C C A A T T T A T T G T G C G T G G T C C T T A C T A C A C G G E T C T A T G T C C C T T

 a - L V T V V K N N K R P R I D V P D T G E -

 13921 - G A A A C C T T G T A T A T C T G A T G C T T G C C A T T T G T G T G T T A T T G T A G T G G T T A T T A T C T T
 C T T T G G A C A T A T G A C T A C G G A C A C G G T T A A A C M A C H A C C A A C C O A A T A T G A A + 13980

 a - E T L Y I L M L V A I L L F G S G Y Y L -

 13981 - A C G A A A A A C C O A A A T T A C T G A T T C A A G T C A T C A T G A M A A G A T A C G G G C T G A
 T G C T T T T T T G G T T T A T T G A C T A T A G T T G A T G A T G A T C T T T C T A T C G T C C G A C T + 14040

 a - T K K P N N *

 14041 - A G G G A A G C C A G A G T C T C A G G T C A T G T T A A T C A G G A A T C A T G G T G A T G T G G C A T G A A + 14100
 T C C C T T T G G T C T A T G A G A C T C C A C T A C A T G T A A T T A G T C C T T A G T A C C A C T A C A C C G T A C T T

Figure 101AH

14101	TCACAAATAACGGATAATGAGGCTGGCAGATTGTGCCAGCTCATTTGGTATTGTTG	14160
	AGTGTATTGGCTATACTCCGACCCGTCTAACACGGTGGAGTAACACCCAATAAAC	
14161	TAACACGATAGGACTGGTCTGGTAATCATTAGGATGGACAGGACTGGATTCTGATT	14220
	ATTTCGCTATCCGACCCATTAGTAAATCCTTACCTGTCTGACCCCTAAGCTAA	
14221	AAAAATGGATGGAAATCAGAAAATGAGATTCTCGTTCTCTAGCAGATAGGAT	14280
	ATTTTACCTTACCACTTAGTCCTTTACTCTAAAGGAGAACGAAATCGCTTATCCTA	
14281	TGTCCTGTTAGGAAAGCGATAAAATGATGAGTTGAAGATAAAGGGATGCTGATAAAAT	14340
	ACAGACAATCCCTTTCGCTTAACTACTCAAACCTCTATTTCCTACGGACTTTTTA	
b	M L I K M -orf6_670, homologue of sp0466, sortase	-
14341	GGTAAAAACAAAAAGCAAAACGAAATAATCCTTATTAGGGGGTATTTCATTTGG	14400
	CCATTTTGTTTTCGTTTTCGTTTATTAGGGATAATCCTCACCCATAAAAGTAACC	
14401	AATGGGGTAATGGGTATCGCGTCTCGCTTGTATTATCGAGTGGAAATCAAAATCA	14460
	TTACCGCCATTACGGCATGGGACCAACAGGGAAACATAATAGCTCACCTTAGTTAGT	
b	V K T K Q R N N L L G V V F F I G -	-
14461	ACAAAATTGCTGACTTTGATAAGGAAAAAGCAAGTGGATGAGGCTGACATTGATGAAACG	14520
	TGTTTAACGACTGAAACTATTCTTTCTGTTGAAATAATGAGCTAACACTACTTGC	
b	Q I A D F D K E K A T L D E A D I D E R -	-
	AATGAAATTGGACAAGCCCTCAATGACnCTTTGAATAATGAGCTGAGTGGCAGTCCTTG	

PCT/US05/27239

286/487

Figure 101AI

b M K L A Q A F N D S L N N V V S G D P W - + 14580
 14581 TTA CTT TAA CCG GTT CGG AA GT TACT GTG AAA ACT TTACATCACTCACCGTAGGAAC
 GTCGGAGAAATGGAGAAAAAGGGCGAGCAGAGTATGCACGTATGTTAGAAATCCATGA
 CAGCCCTCTTTACTTCCTTTTCGGCTCGTCTCATACGTGCATACAATCTTGGTACT + 14640

b S E M K K G R A E Y A R M L E I H E - + 14641
 GCGGATGGGGCATGGAAATCCCCGTATTGACGTGGATTGCCGGTTATGCCGTAC
 CGCCTACCCGTAACCTTTAGGGCAATAACTGCACCTAACGGCAAATACGACCATG + 14700

b R M G H V E I P V I D V D L P V Y A G T - + 14701
 TGCTGAAGAGGTATTGCAGCAAGGGCTGGGCATCTAGGGAAACTTCTCTGCCGATCGG
 ACCACTTCCTCCATAACGTCTGGTCCCCGACCCGTAGATCTCCCTTGAGAGACGGCTAGCC + 14760

b A E E V L Q Q G A G H L E G T S L P I G - + 14761
 AGGCAATTGACCCATGGGTGATTACGGCACATACAGGTGTTGCCAACAGCTTAAGATGTT
 TCCGTTAACGCTGGTACGCCACTAATGCCGTGTTATGTCGATTCACAA + 14820

b G N S T H A V I T A H T G L P T A K M F - + 14821
 TAGGGATTGACCAAACCTTAAAGTGGGATAAGTTTATGTGCACAAATCAAGGAAGT
 ATGCCTAACGGTTGAAATTCAACCCCTATTCAAACAGGTTGCTTATAGTTCCCTCA + 14880

b T D L T K L K V G D K F Y V H N I K E V - + 14881
 GATGGCTTATCAAGTGGATCAAGTAAAGGTGATTGAGCCGACGAACTTTGATGATTATT
 CTACCGGATAGTTCACCTAGTCACTTCAACTAACTCGGCTGCTGAAACTACTAAATAA + 14940

b M A Y Q V D Q V K V I E P T N F D D L L -

PCT/US105/22239

287/487

Figure 101AU

PCT/US05/27239

288/487

Figure 101AK

C S N E V I K E F D E T V S Q M D K A E L -
 AGTCAAACGAGGTATTAAAGAGTTTGATGAGACGGTTCCAGATGGATAAGCGAAC
 15361 TCAGTTCGCTCAATAATTCTCAAACACTCTGCCTAACCTATTCCGTCTTG
 15420

C E E R W R L A Q A F N A T L K P S E I L -
 TTGAGGAGCCTGGCCTTGGCTCAAGCCTTCAATGGACCTTGAACCATTGAAATTC
 15421 AACTCTCGCAACCGGAACCGAGTTGGAAAGTTACGGTGGAAACTTGGTAGACTTTAAG
 15480

C D P F T E Q E K K G V S E Y A N M L K -
 AGGTCCATGAGGGATTGGCTATGTGGAAATTCTGCATTGATCAGGAATTCCGATGT
 15481 AACTAGGAAATGTCTCGTTCTCTTCTTCTTCGGAGAGTCTTACGGTTATACGGATT
 15540

C V H E R I G Y V E I P A I D Q E I P M Y -
 ATGTCGGAACGAGGTGGAAATTCTTCAGGAGGGCGAGGATTGCTAGAGGGGCTTCGT
 15541 TACAGCCTTGCTCACTCCTTAAGAAGTCTCCGGCTCCTAACGATCTCCCTCGAAGCA
 15600

C V G T S E E I L Q K G A G L L E G A S L -
 TACCGGTTGGTGGAAATTACCCACACAGTTGTCACGTGCTCATAGGGATTACCGACGG
 15661 ATGGCCAACCAACTTTATGGCTGTCACAGTGACGAGTATCTCCCTAATGGCTGCC
 15720

C P V G G E N T H T V V T A H R G L P T A -
 CAGAAACTGTTAGTCAATTGGATAAGATGAAAAAGGGGATGTCTTATCTTCAGGTT
 15721 GTCTTGACAATTAGTTAACCTTACTTCTACTTCTACGAAATAGAAGTGGCAA
 15780

PCT/US05/27239

289/487

Figure 10A

c E L F S Q L D K M K G D V F Y L H V L -
 TAGACCAGGTGTTGGCCTACCAACTGATCAGATTGACGGTGTAGCCAAATGGACTTTG
 15781 ATCTGTCACCAACCGGAACTTCACCTAGCTAAAACCTGCAACTGGTTTACTGAAAC

c D Q V L A Y Q V D Q I L T V E P N D F E -
 AGCCTGTTGATTCAACATGGGAGATTATGCCACCTTGTGACTCTGTACACCGTATA
 15841 TCGGACAGAACTTAATTTGATCCCCCTCATATACCTGGAAACAACCTGACATGTGGCATAT

c P V L I Q H G E D Y A T L L T C T P Y M -
 TGATTAAACAGTCATCGTCTGTGTGTGTACGTGGGAAGGGATTACGGCACCAAATTG
 15901 ACTAAATTGTCAGTAGCACACAACTGACCCATGGCCATATGCCATATGCCGTGGTTAAC

c I N S H R L L V R G K R I P Y T A P I A -
 CAGAGCCAAATCGGGGTCAAGAGAGGTTGGGCAATTCTGGTTGTGTGTATTGCTAGGG
 15961 GTCTCGCTTAACTCGCCACTCTCGCACCGTTAACACCAACAAATAACGATCGCC

c E R N R A V R E R G Q F W L W L L A A -
 CGTTGGTTATGATTCTGGTATTGAGTTACGGGGTGTATCGTCATCGCATTTGTCAAAG
 16021 GCAACCAAATCAACTAAAGACCAATCAATGCCAACATAGGCACTAGGACGGTAACAGTTTC

c L V M I L V L S Y G V Y R H R R I V K G -
 GGCTTAAGAAACAAATTGGAGGAGCATATGCTAAAGGCTAAGGTACAGAAATTACTAGGG
 16081 CCCATCGTTTGTTGTTAACCTCCCTCGTACAGTTCCGATTGATGTCCTTAATGATCCC

a M S K A K L Q K L L G -orf8_670, homologue of sp0468, sortase

c L E K Q L E H H V K G *

PCT/US05/27239

290/487

Figure 101AM

Figure 101AM

<pre> TATTTGGCTAATGCTGGTAGCAATTGGGAAATTCCCGTTAATGTTGGGAGATGGGTAA ATTAACGACTAACCCATCGTAACCACTAAGGACAATAACAAAACCCGTCTACACAAAT </pre>	16141
<pre> CAGTCTTCTAGGACAAGTAAAGGTCAATGAGATTTCTAGAAATCTGTGACGGCCGACAGT GTCAGAAATCCTGTCTCATTTCCAGTACTCTATAAAGCTTCTAGACACTGCCGGCTGTCA </pre>	16201
<pre> a Q S L G Q V K G H E I F S E S V T A D S - TACCOAGGAAATTGCAACGGTGGCTTGATACAACTAACGGCTGGATTGCAAATGCT </pre>	16261
<pre> ATGGTTCTCGTTAACGGTGGCAAGGAACTAATGTTAGTTAGCTTGGAACTTAAGGGTTTAGCA </pre>	
<pre> a Y Q E Q L Q R S L D Y N Q R L D S Q N R - ATTGTAGATCCTTTTGGGGAAAGGTATGAGGTAAATTACCAAGTGTCTGACGATCCT </pre>	16321
<pre> TAACATCTAGGAAAAACGGCTTCCCATCTCTCCATTAACTGTTACAGACTGCTAGGA </pre>	
<pre> I V D P F L A E G Y E V N Y Q V S D D P - GATGGCAGTCAACGGCTATTGTGGATTCCGAGTTGGAAATCATGGAGCCAGTTATCTA </pre>	16381
<pre> CTACGTCAGATGCCGATAAACGCTAACGGCTAAACCTTATGACTCTCGGTCAAATAGAT </pre>	
<pre> a D A V Y G Y L S I P S L E I M E P V Y L - GGAGGGGATTACCATCATTAGCAATGGGGTTGGCCCATGTGGATGGGACGCCTCTTCTC </pre>	16441
<pre> CCTCGCTTAATGGTAGTAAATGTTACCCCAACGGGTACACCTACCCCTGGGAGAAGGA </pre>	
<pre> a G A D Y H H L A M G L A H V D G T P L P - GTTGAGGGAAAGGATTCTGTTCACTGTTAGTGTGGGACCGTGGAGAACCAAGCCTGTC </pre>	16501
<pre> CAACTCCCTTTCCCTAACGAAAGTCACTAACGACCCGGTGGCACETCTGGTACAG </pre>	

Figure 101AM

a V E G K G I R S V I A G H R A E P S H V -

TTTTTCGCCATTGGATCAGCTAAAGTGGAGATGGCTCTTTATTATGATAATGGCCAG
16561 AAAAGGGGGTAAACCTAGTCGATTTCACCGAGAATAACTATTACCGGTC
16620

a F F R H L D Q L K V G D A L Y Y D N G Q -

GAAATTGTAGAATATCAGATGATGGACACAGAGATTATTTACCGTGGAAATGGAAAAA
16621 CTTAACATCTTATAGTCTACTACCTGTTCTAAATAATGGCAGGCCCTAACCTTTT
16680

a E I V E Y Q M M D T E I I L P S E W E K -

TITAGAAATCGGTTAGCTCTAAATAATCATGACCTTGATAACCTGCGATCCGATTCCTTAC
16681 AATCTTAGCCAATCGAGATTATAGTACTGGAACTATTGGACGCTAGGCTAAGGATGG
16740

a L E S V S S K N I M T L I T C D P I P T -

TTTAATAAACGCTTATTAGTGAATTGGAAAGCAGTCGCTGTTTATCAAAAATCAGATCCA
16741 AAATTATTTGCGAATAATCACTAAACTTGCCTAGCGACAATAAGTTTCTAGCTAGGT
16800

a F N K R L L V N F E R V A V Y Q K S D P -

CAAACAGCTGCAGTTGCAGGGTTGCTTACGAAAGAAGGACAATCTGTATCGCGTGT
16801 GTTTGTGACGTCAAGCTCCAAACGAAAATGCTTCTCTGTAGACATAGGCCACAA
16860

a Q T A A V A R V A F T K E G Q S V S R V -

GCAACCTCTCAATGGTTGTACCGTGGCTAGTGGTACTGGCATTTCTGGGAATCCTCTGGTT
16861 CGTTGGAGAGTTACCAACATGGCACCGATCACCAGCCGTAAAGACCCCTAGGACAAA
16920

a A T S Q W L Y R G L V V I A F L G I L F -

GTTTTGGAAAGCTAGCACGGTTACTACGAGGGAAATAAAAALGAAGGAAAGCTA
16921

16980

PCT/US05/22239 292/487

Figure 101AN

CAAAACACCTTCGATCAGCCAAATGATCCTCCCTTATTTCCTTTACTTTCTTCGAT

a V L W K L A R I L R G K *

16981 AGGGTGTTCCTTTCGGGCTCTTCTGAACTCTAGTGGTTGAAAAAAGCTAAGCTCG
 TCCGACAGGAAAAGGCCAGAACAGTTGACATCACCAACTTTTGATTCGAGC

17040 AGAAAGGACAATAATTGTCCTTCCTTGTATTCAGAGCGATAAAATCCGGTTTT
 TCTTTCCTGTTAAACAGGAAGAAAAACTATAAGTCTCGGTATTTTGGCAAAAAA

17101 GAAGTTTCAA
 17112 CTTCAAAAGTT

PCT/US05/27239

293/487

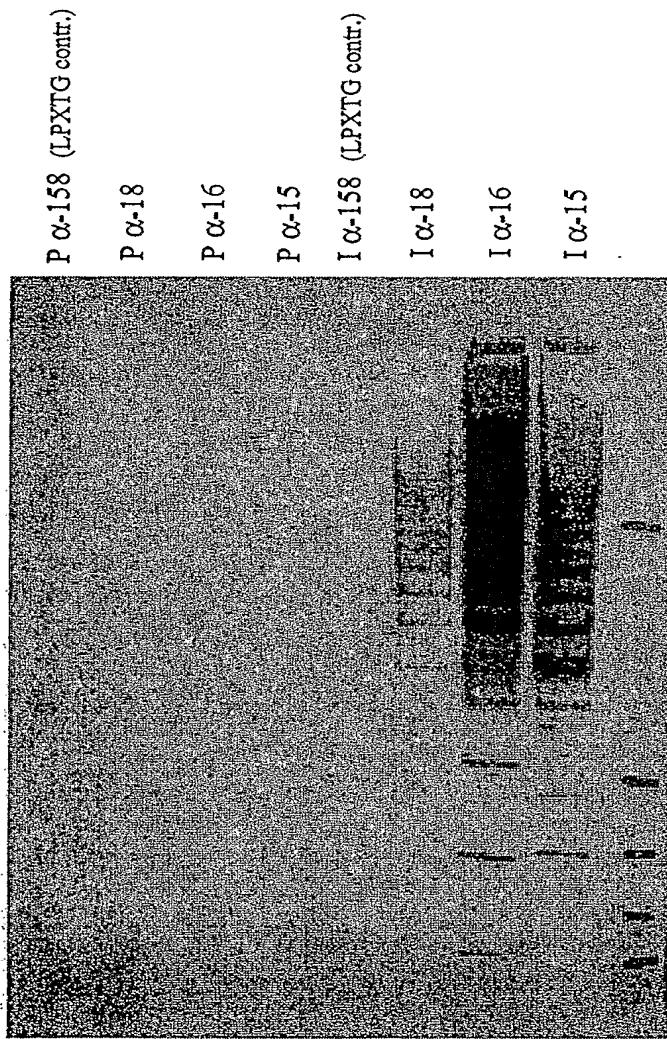
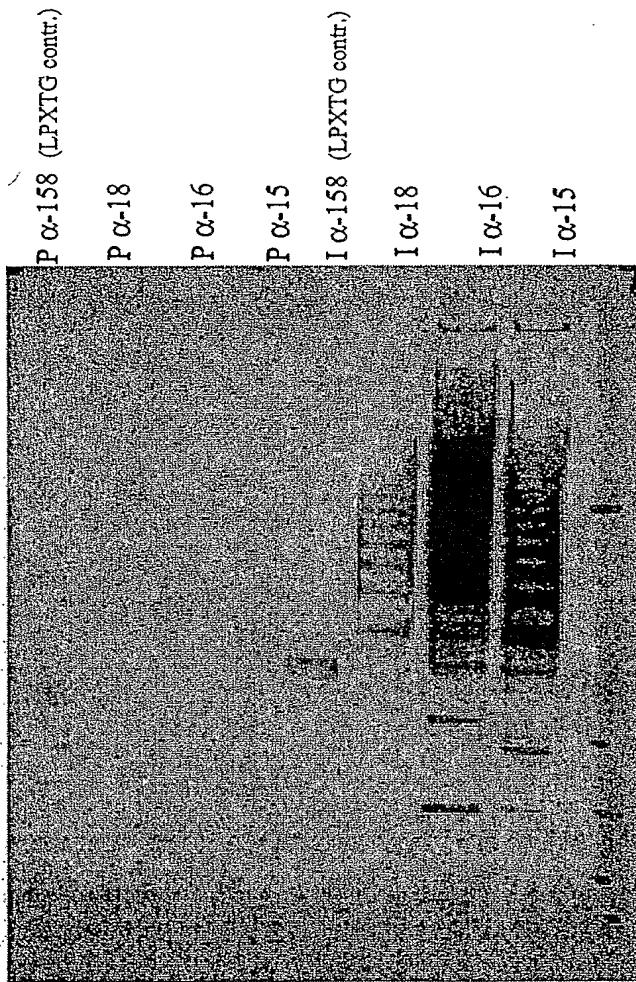
M1, strain 2580

Figure 102

LEGEND:

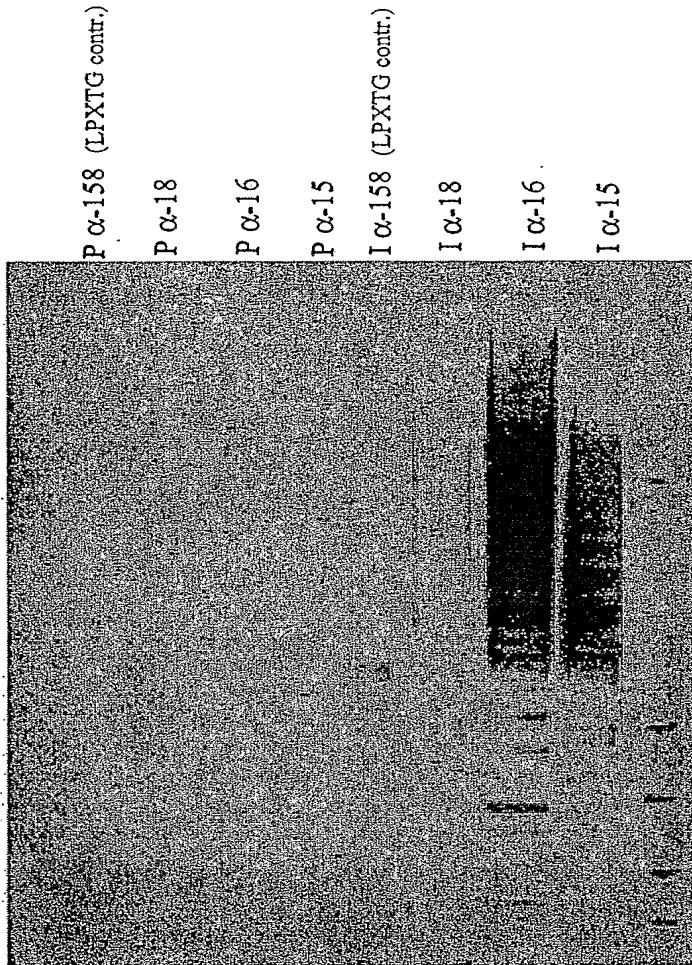
I α -#: immune serum anti-#
P α -#: pre-immune serum anti-#

PCT/US05/27239 294/487

M1, strain 2913**Figure 103****LEGEND:**

I α -#: immune serum anti-#
P α -#: pre-immune serum anti-#

M1, strain 3280

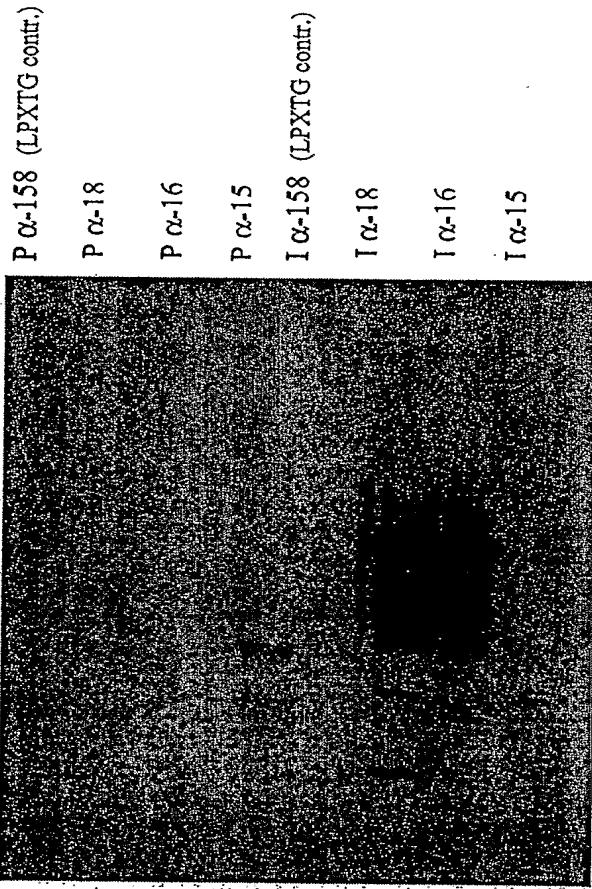


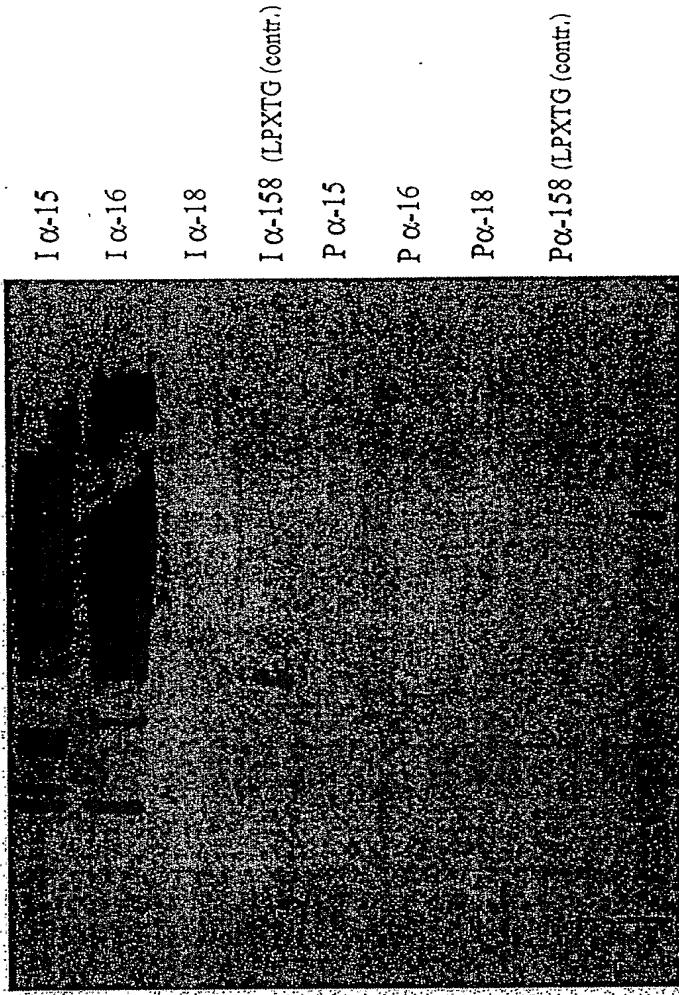
LEGEND:

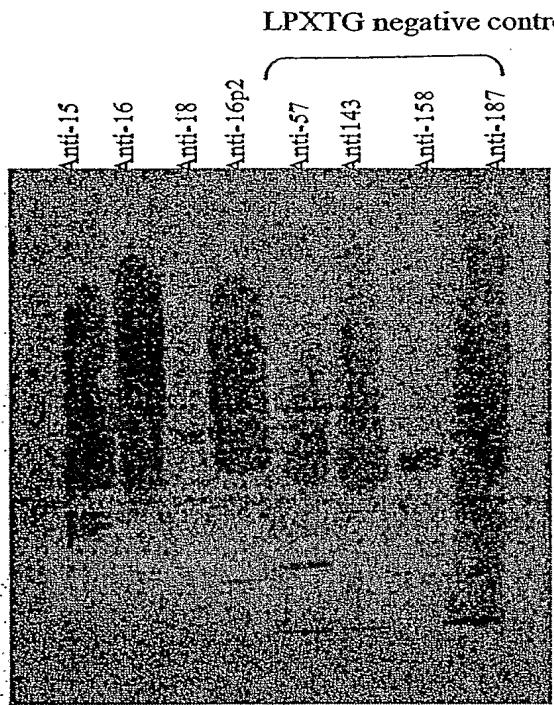
I α -#: immune serum anti-#
P α -#: pre-immune serum anti-#

Figure 104

296/487

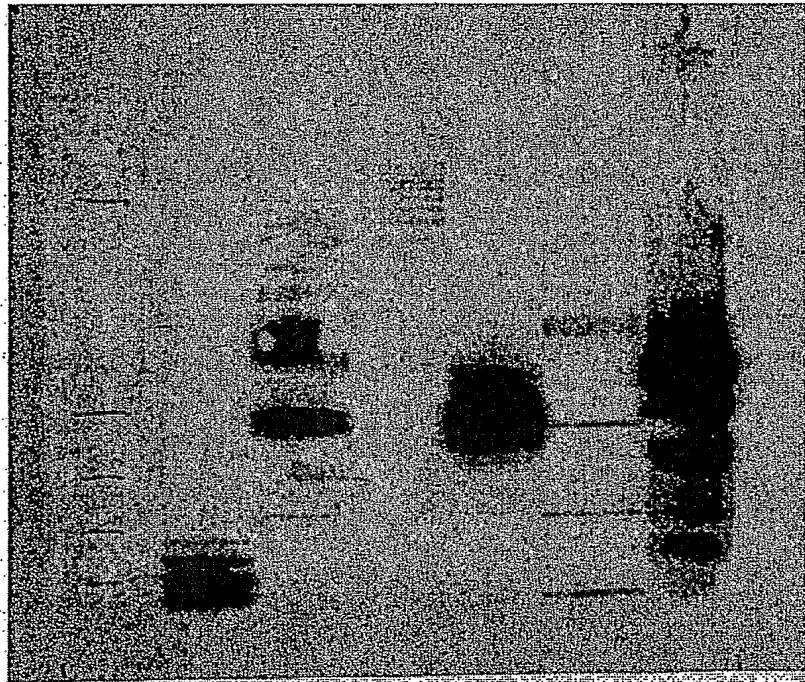
M1 strain 3348**Figure 105**

M1 strain 2719**Figure 106**

Western blot Western blot on fraction enriched in surface proteins of M1 (SF370)**Figure 107**

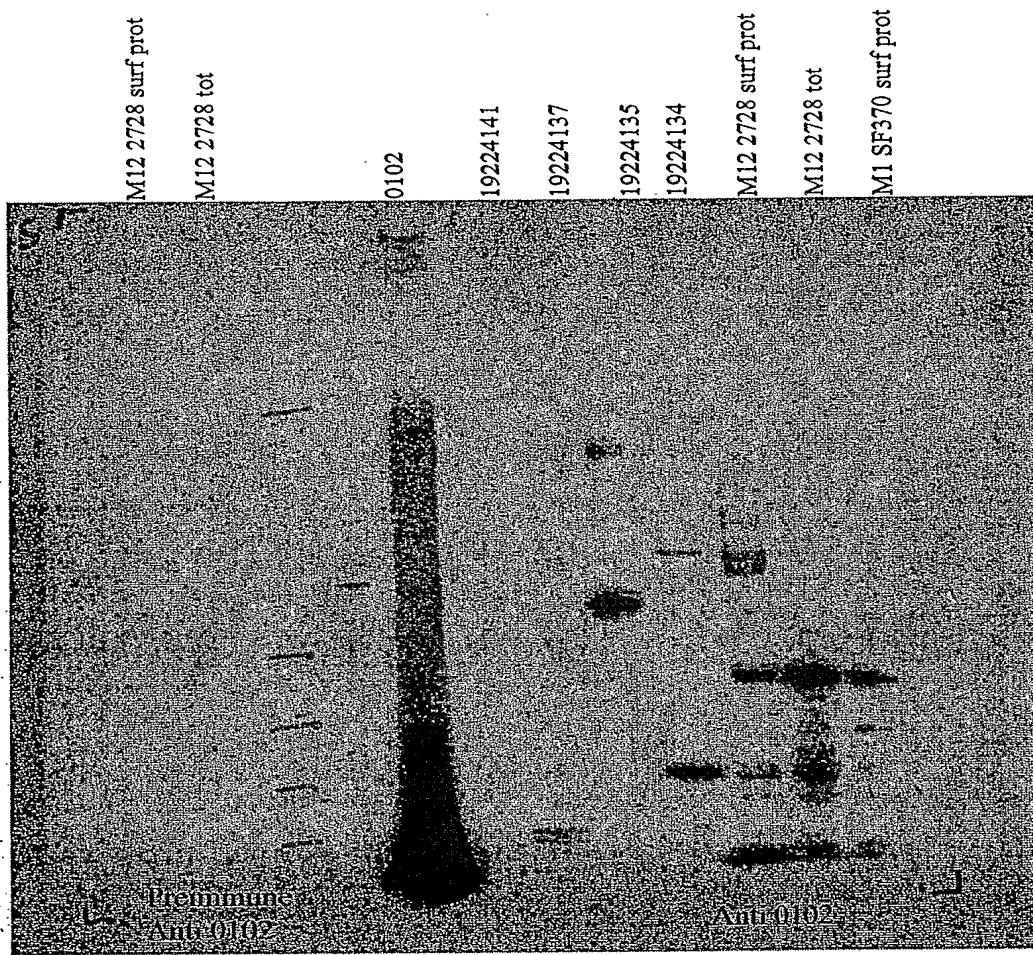
Western blot on fraction enriched in surface proteins of M12 (2728)

Anti-19224134
Anti-19224135
Anti-19224137
Anti-19224141
Anti-0102
Anti-158 (control)



Note: 0102 corresponds to SpyM3_0102 in MGAS315 (M3). This protein is pretty much identical to the M12 19224139 -see also gel "WB M12 (23-11-4b2a)"-, the M18 spyM18_0130 and the M5 orf 82. Here the antiserum raised against 0102 recognizes high-weight molecular structures in M12 (strain isolate 2728). This means that in M12 three out of the five LPKTG proteins form high-molecular weight structures:

Figure 108



Note: 0102 corresponds to SpyM3_0102 in MGAS315 (M3). This protein is pretty much identical to the M12 19224139 -see also gel "WB M12 (23-11-4b2bis)"-, the M18 spyM18_0130 and the M5 orf 82. Here the antiserum raised against 0102 recognizes high-weight molecular structures in protein extracts from M12 (strain isolate 2728.)

Figure 109

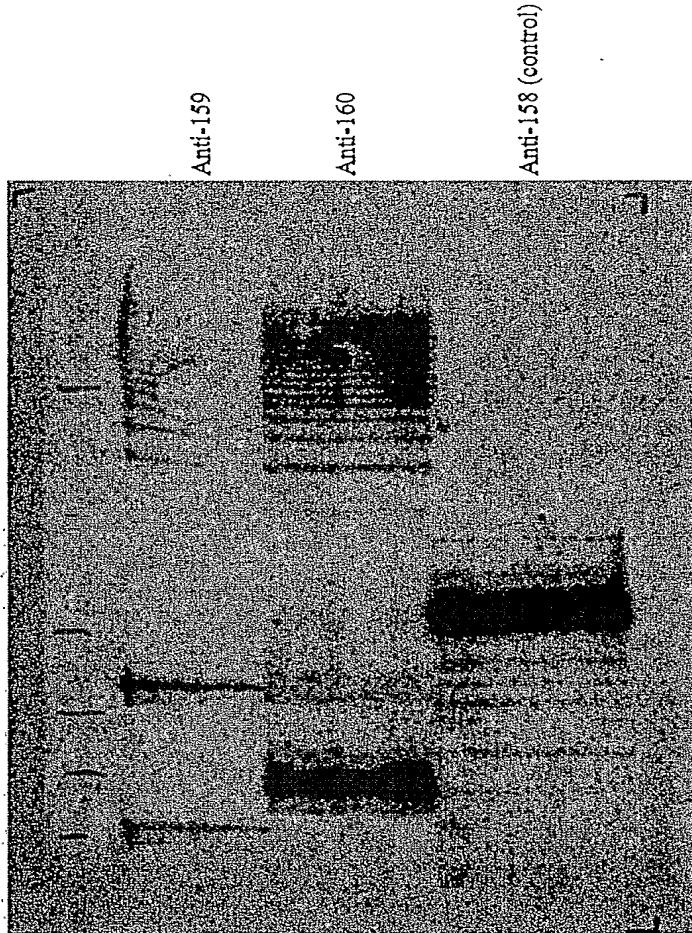
WO 2006/078318

301/487

PCT/US2005/027239

PCT/US05/27239

Western blot on fraction enriched in surface proteins of M6 (2724)



M6 strain isolate 2724

Figure 110

WO 2006/078318

302/487

PCT/US2005/027239

PCT/US05/27239

0.160

0.159

M6 3650 surf prot

M1 SF370 surf prot

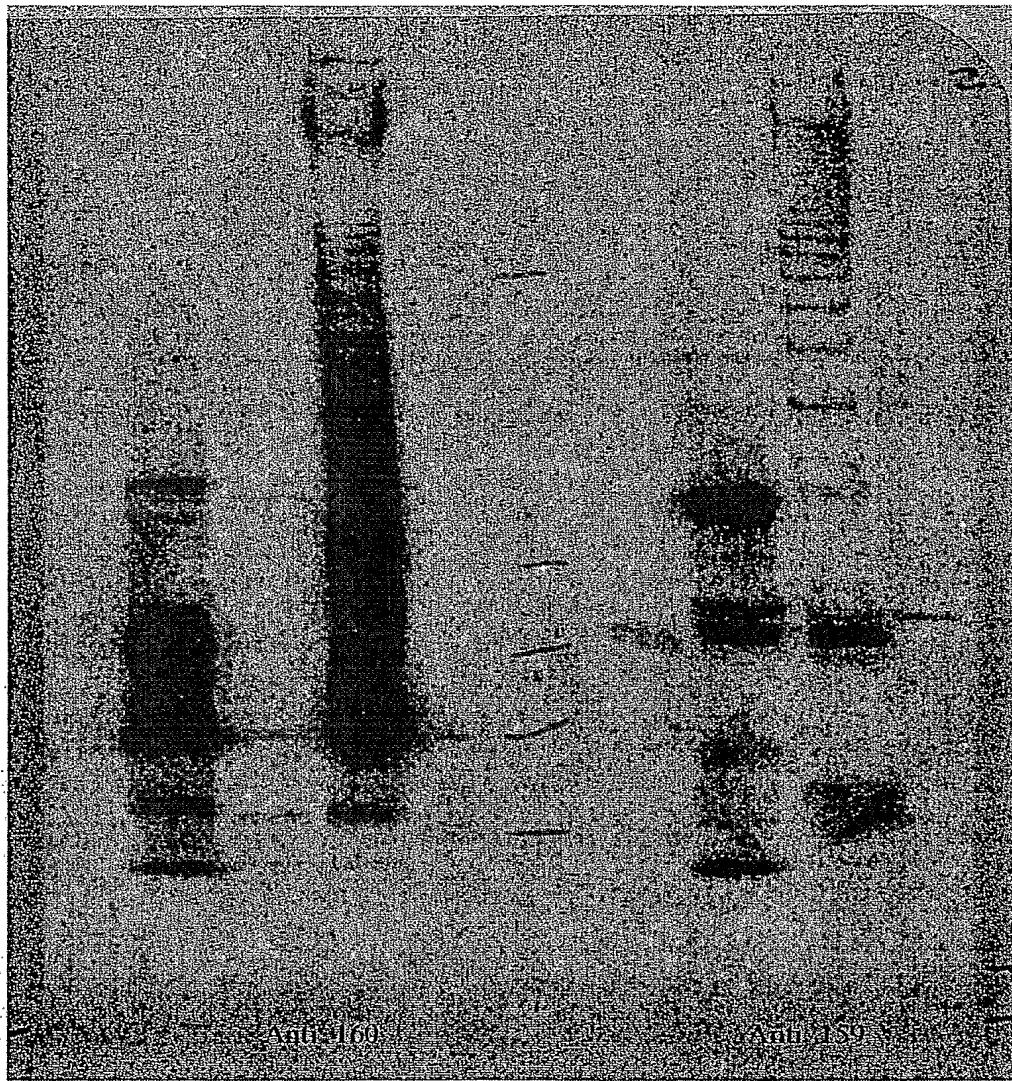
Marker

0.160

0.159

M6 3650 surf prot

M1 SF370 surf prot



M6 strain isolate 3650

Figure 111

WO 2006/078318

303/487

PCT/US2005/027239

PCT/US05/27239

M6 2724 fraz art

M6 2724 tot

0160

Marker

158

0160

0159

M6 2724 surf prot

M6 2724 tot

M1 SF370 surf prot

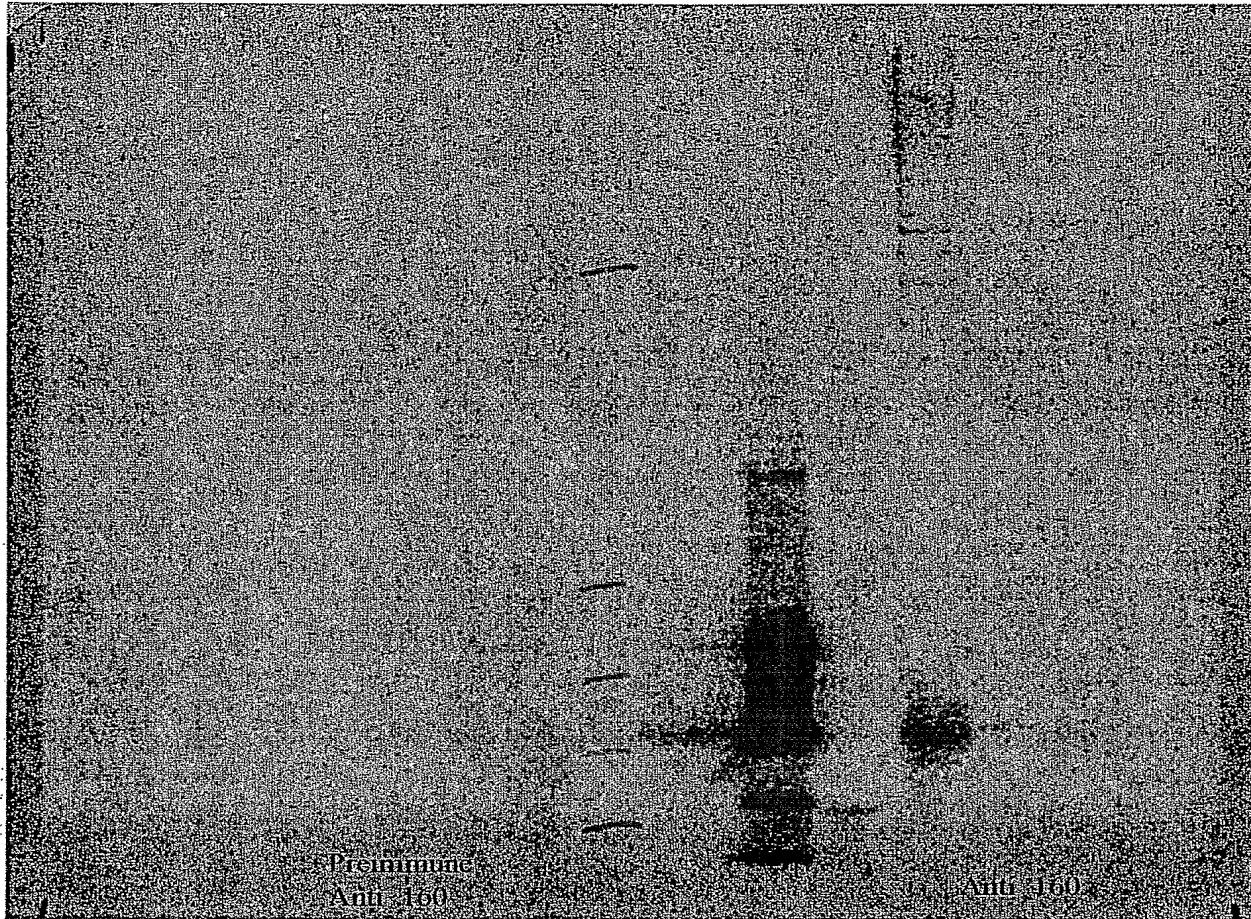


Figure 112

WO 2006/078318

PCT/US05/27239

304/487

PCT/US2005/027239

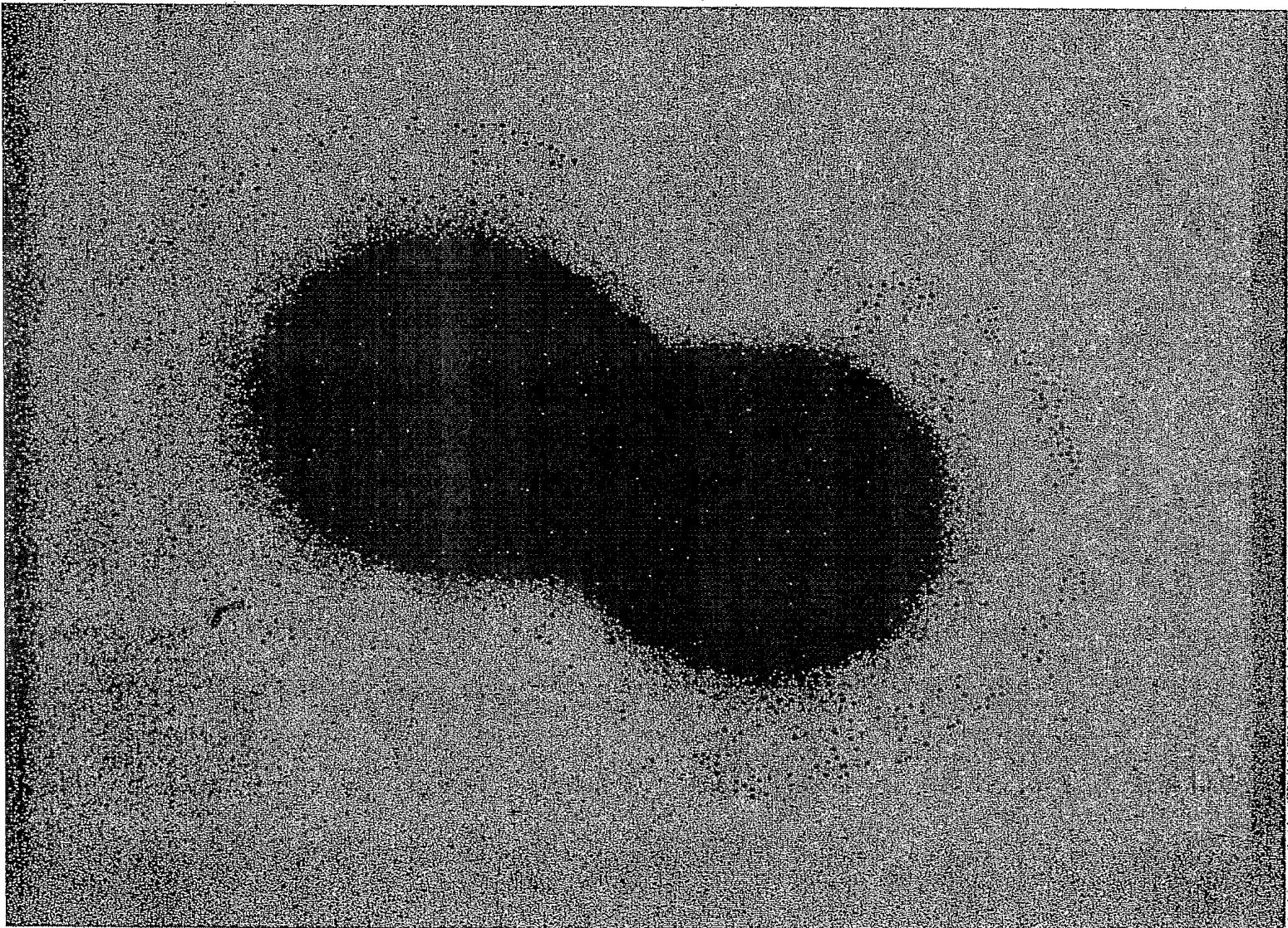


FIGURE 113

WO 2006/078318

PCT/US2005/27239

305/487

PCT/US2005/027239

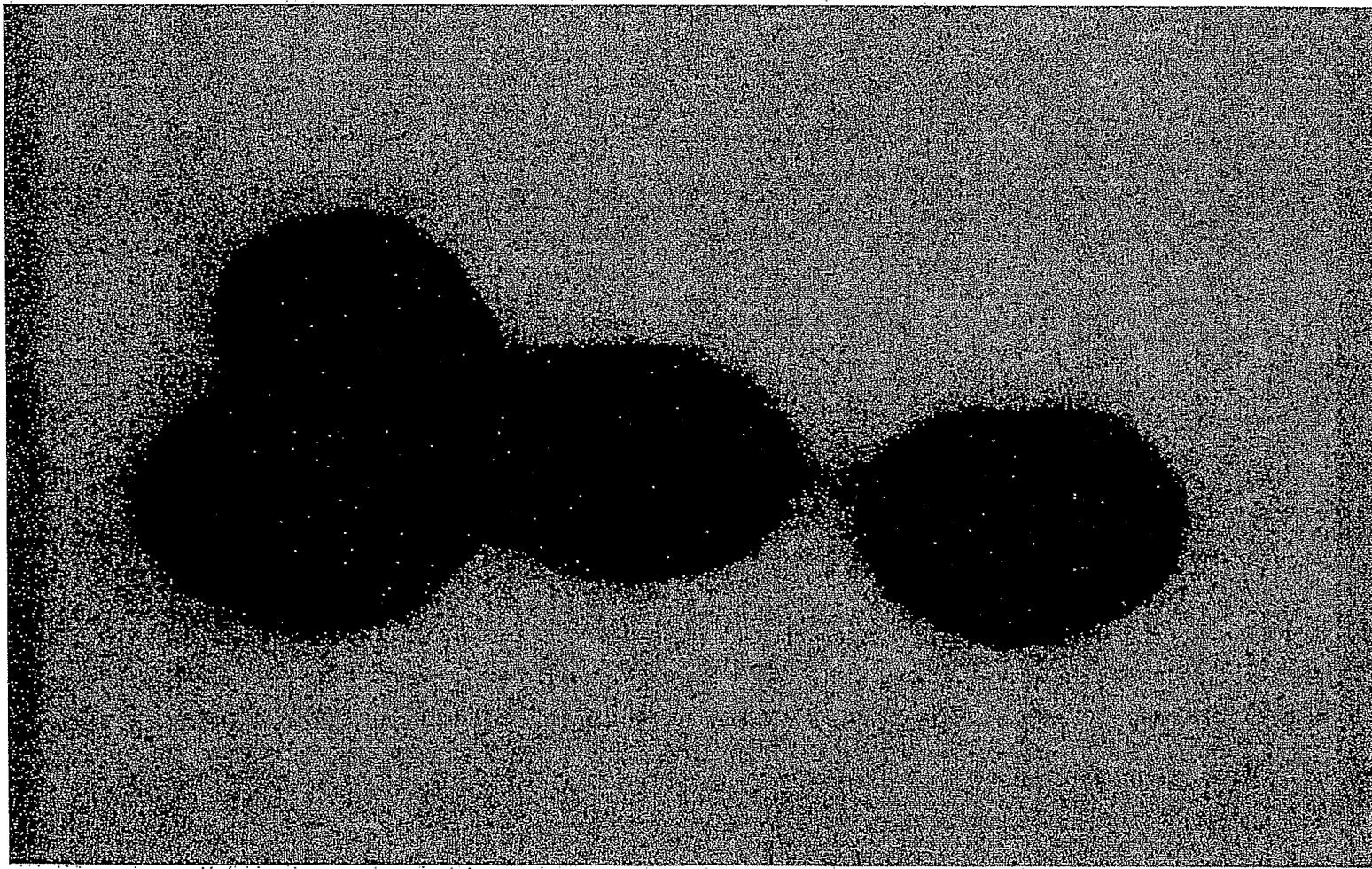


FIGURE 114

WO 2006/078318

PCT/US2005/27239 306/487

PCT/US2005/027239

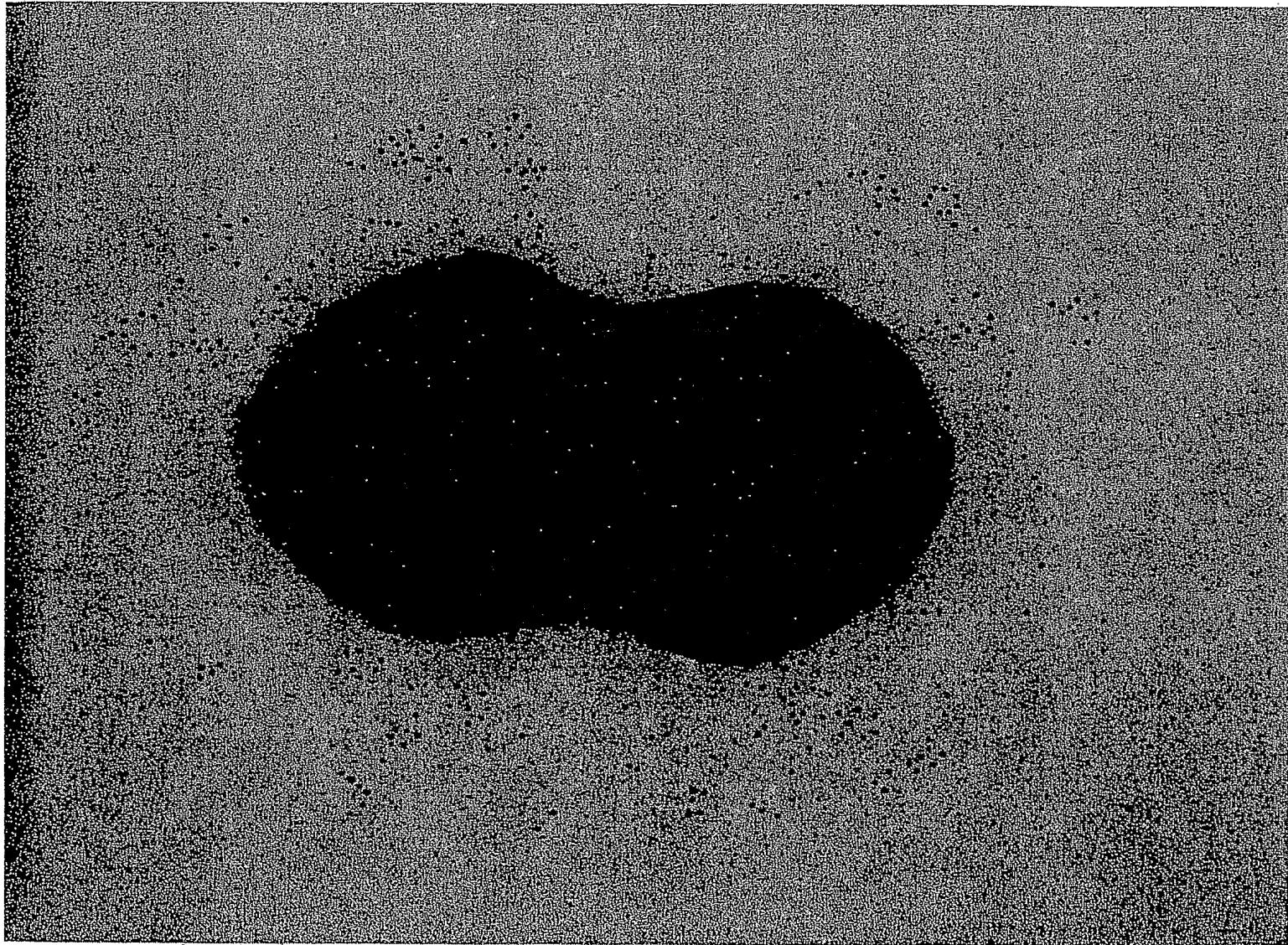


FIGURE 115

WO 2006/078318

PCT/US2005/027239

PCT/US05/27239
307/487

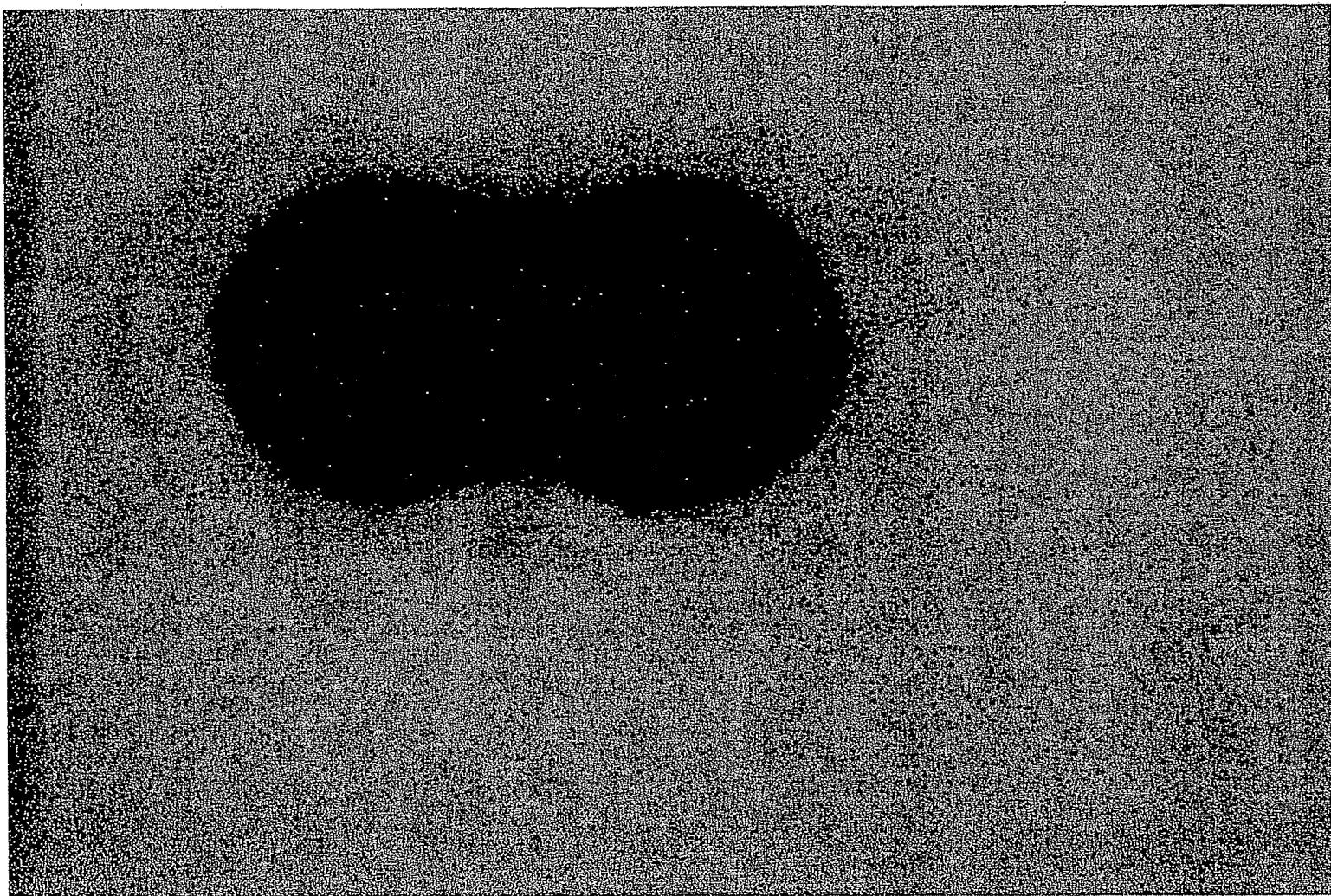


FIGURE 116

WO 2006/078318

PCT/US2005/27239

PCT/US2005/027239

308/487

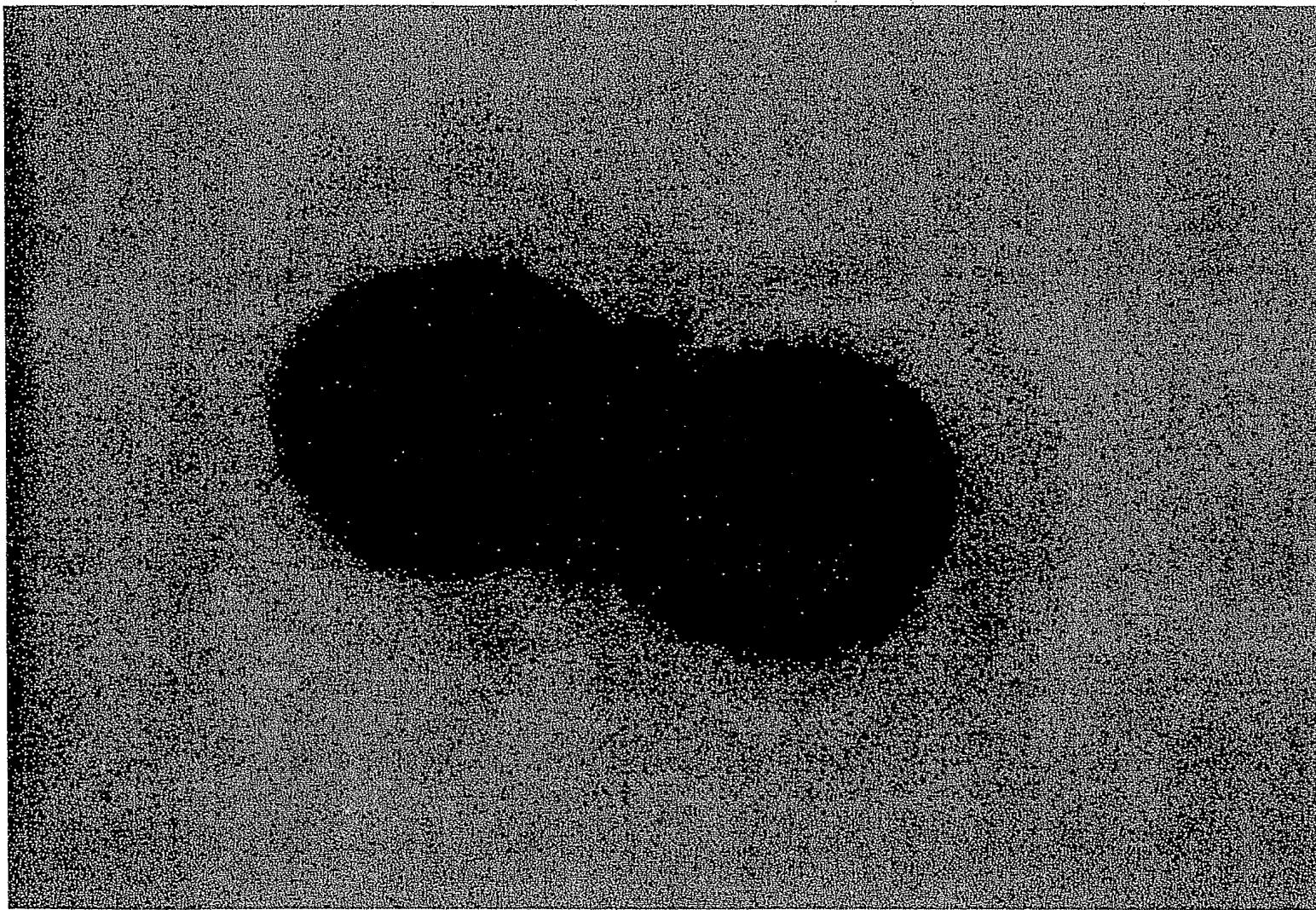


FIGURE 117

WO 2006/078318

PCT/US2005/27239

309/487

PCT/US2005/027239

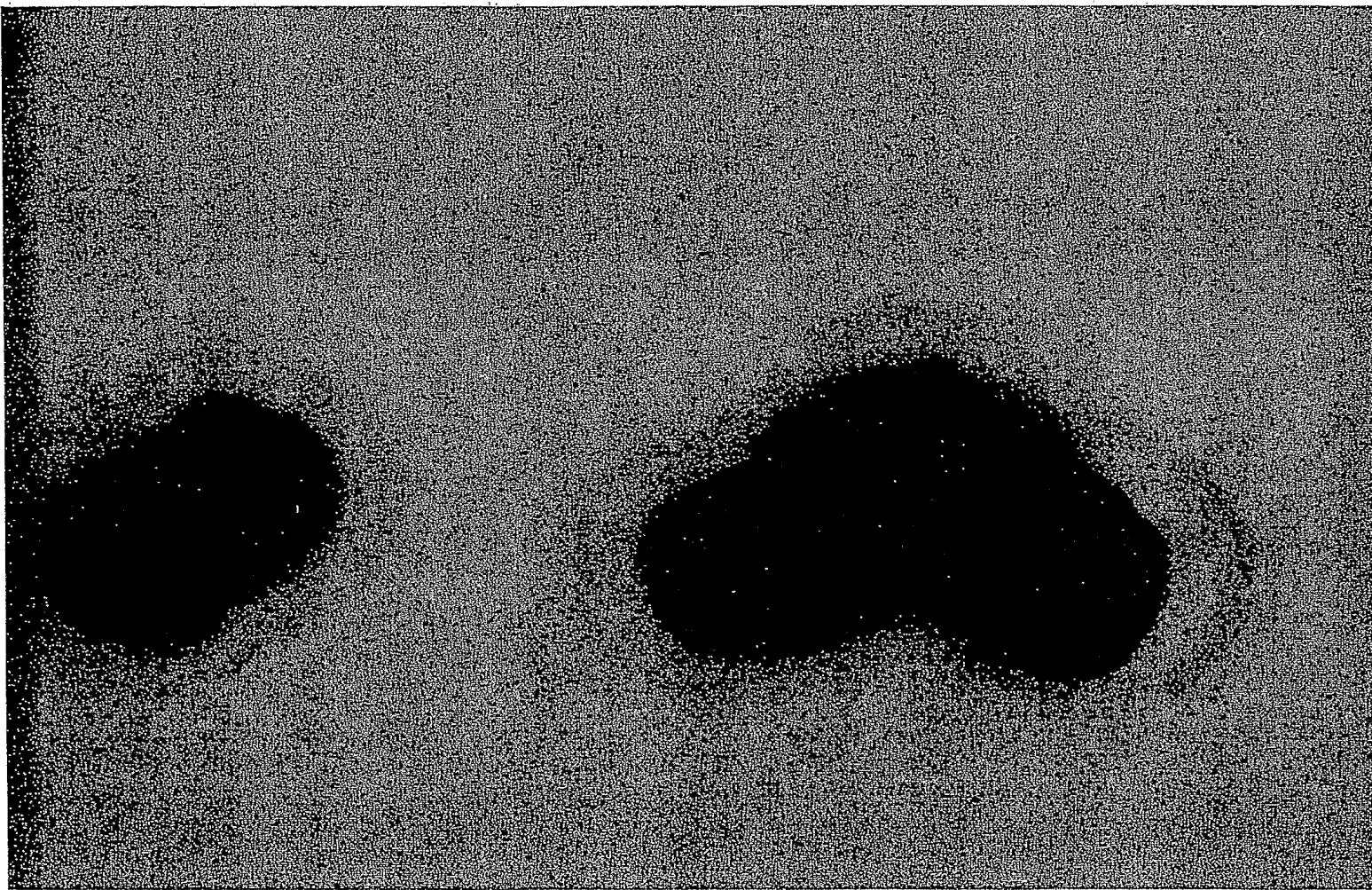


FIGURE 118

WO 2006/078318

PCT/US2005/027239

PCT/US05/27239
310/487

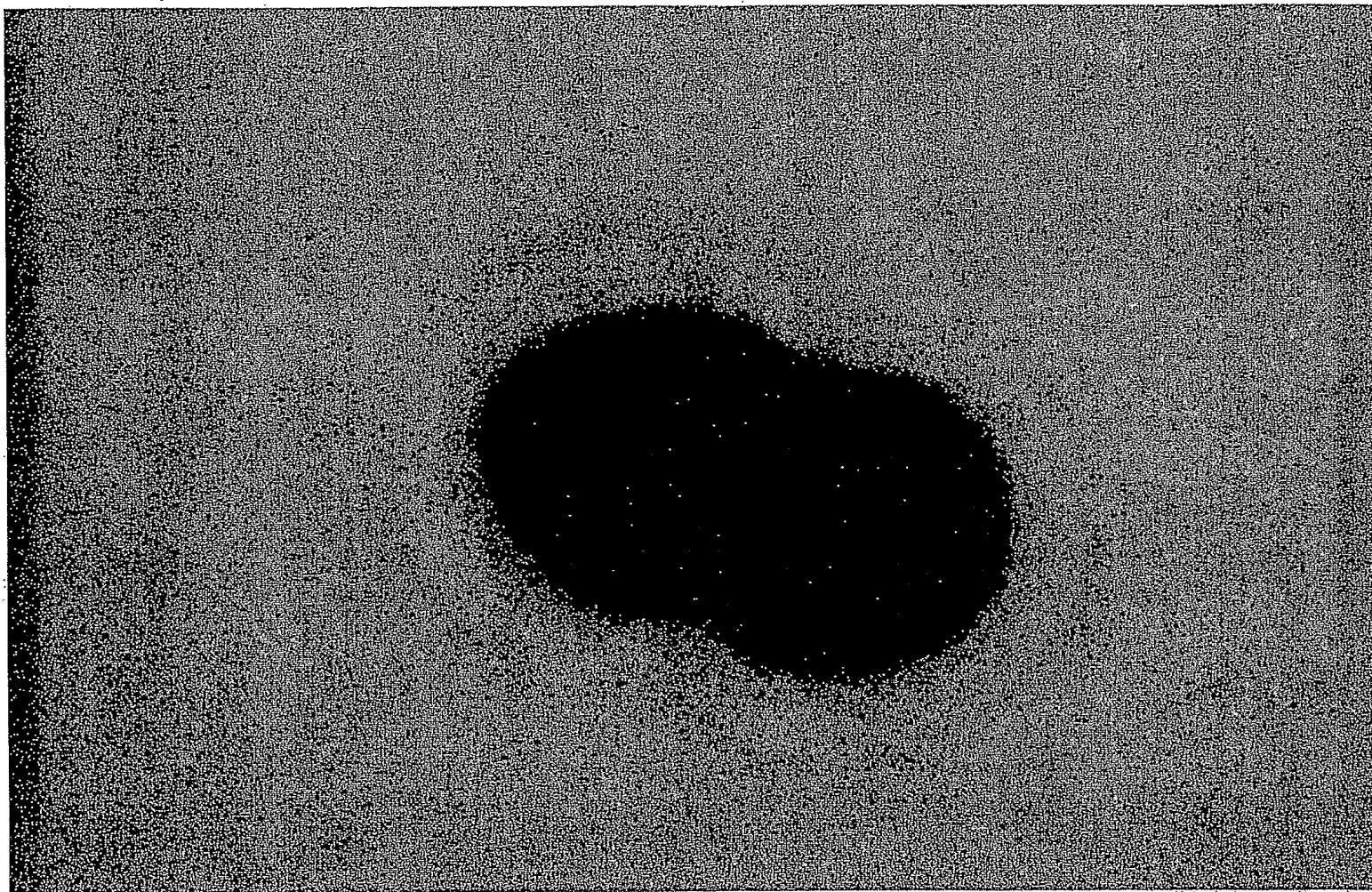


FIGURE 119

WO 2006/078318

PCT/US05/27239

PCT/US2005/027239

311/487

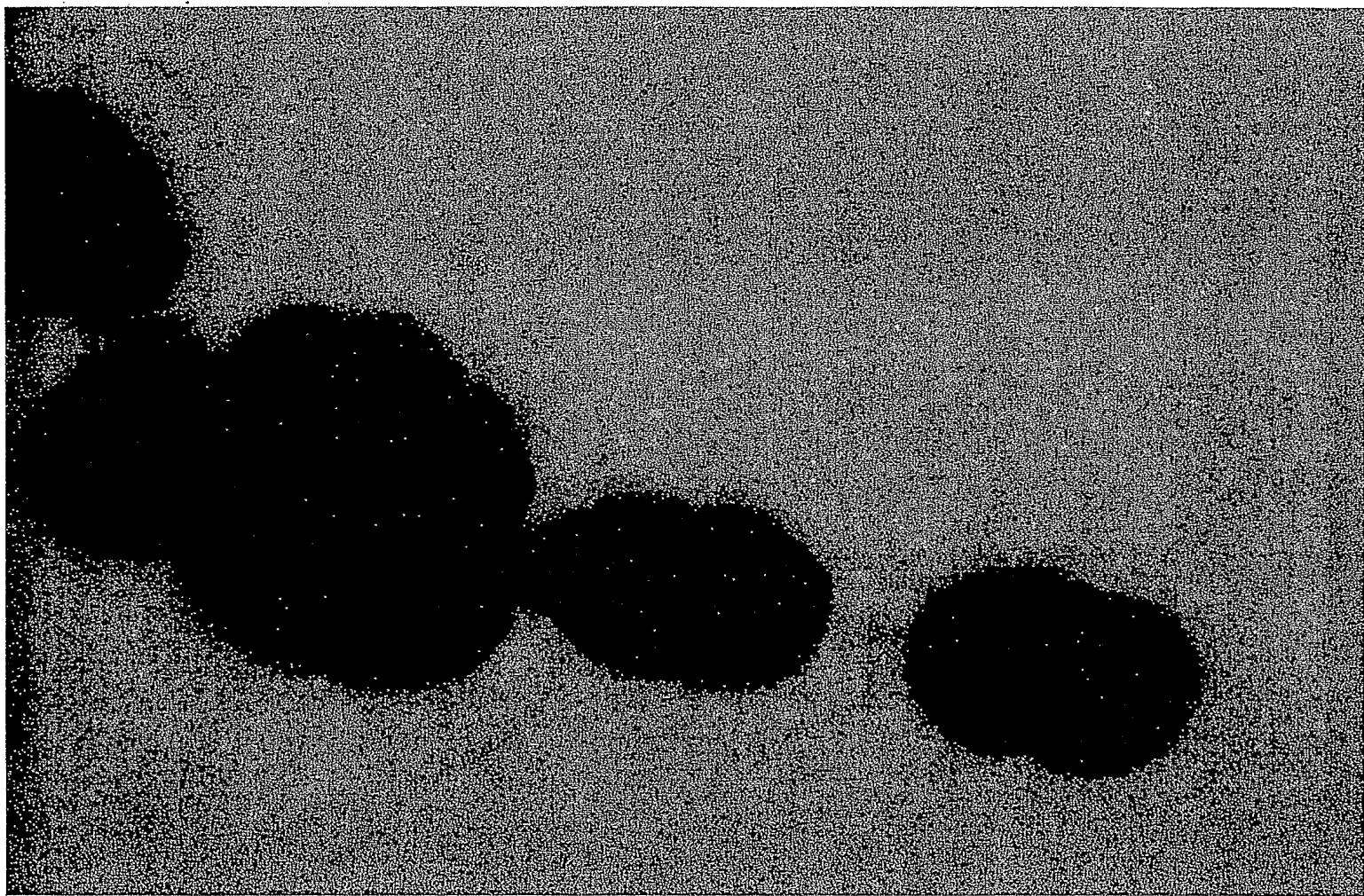


FIGURE 120

WO 2006/078318

PCT/US05/27239

PCT/US2005/027239

312/487

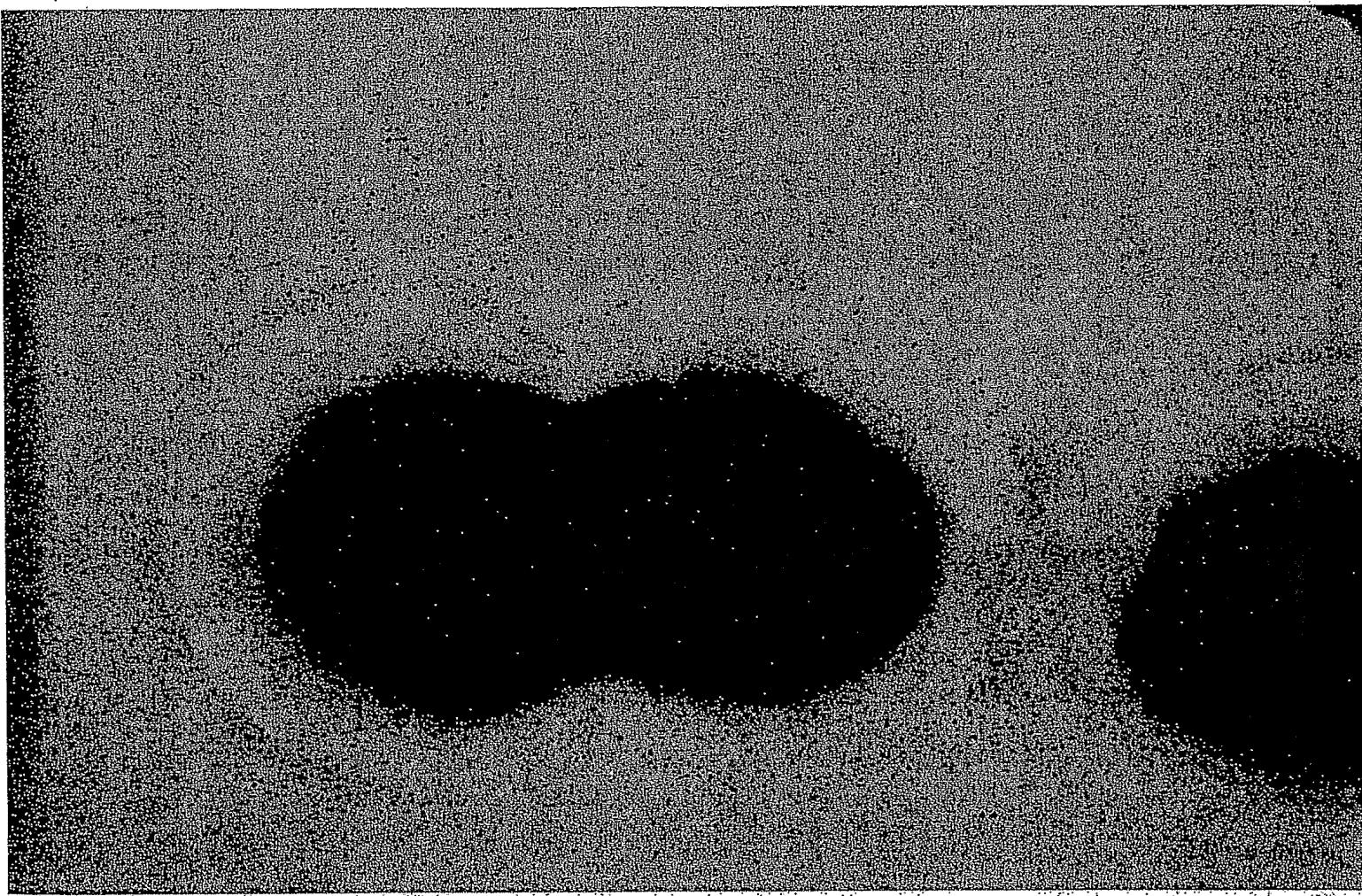


FIGURE 121

WO 2006/078318

PCT/US2005/027239

PCT/US605/27239

313/487

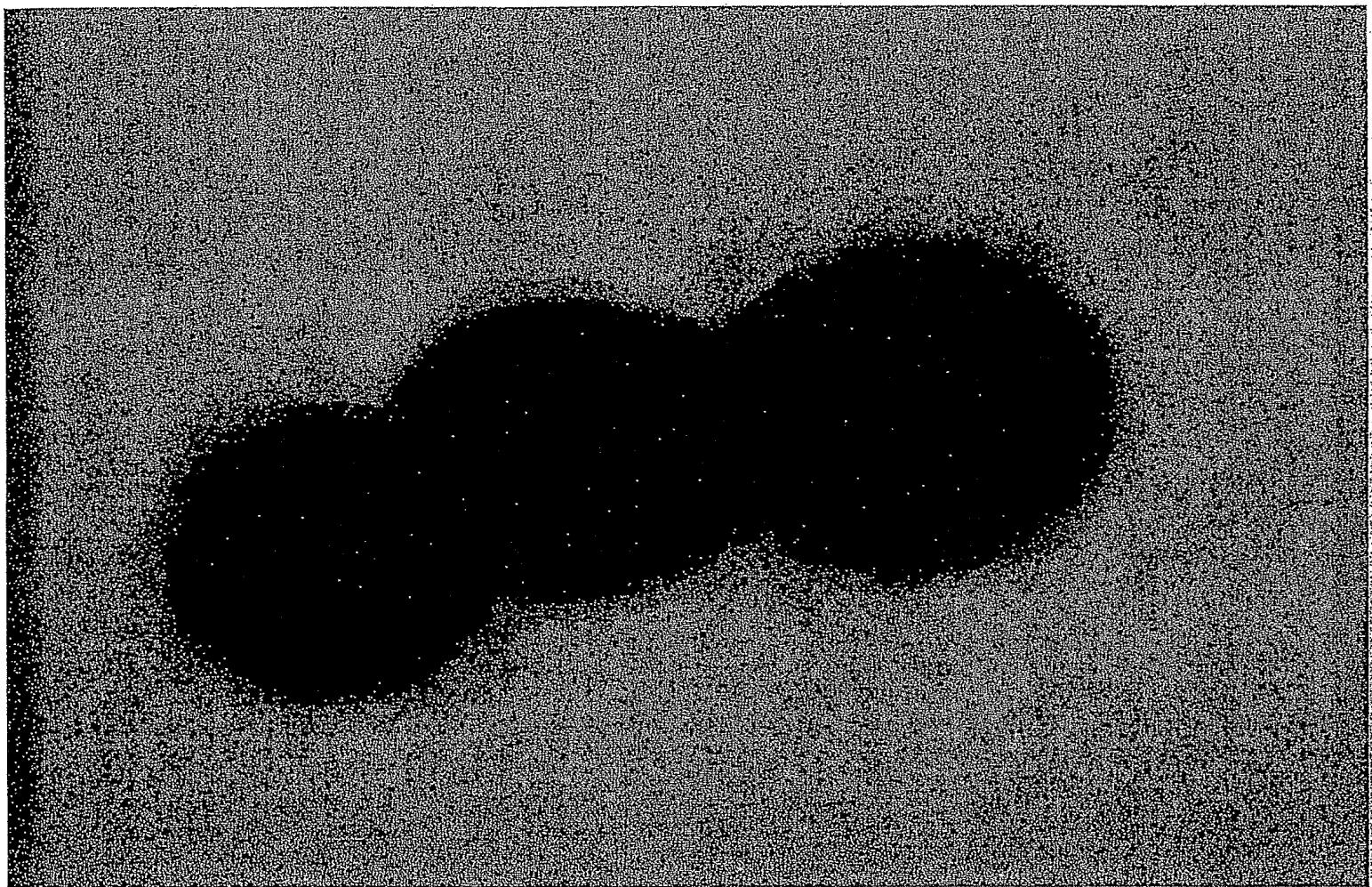


FIGURE 122

WO 2006/078318

PCT/US2005/027239

PCT/US2005/027239

314/487

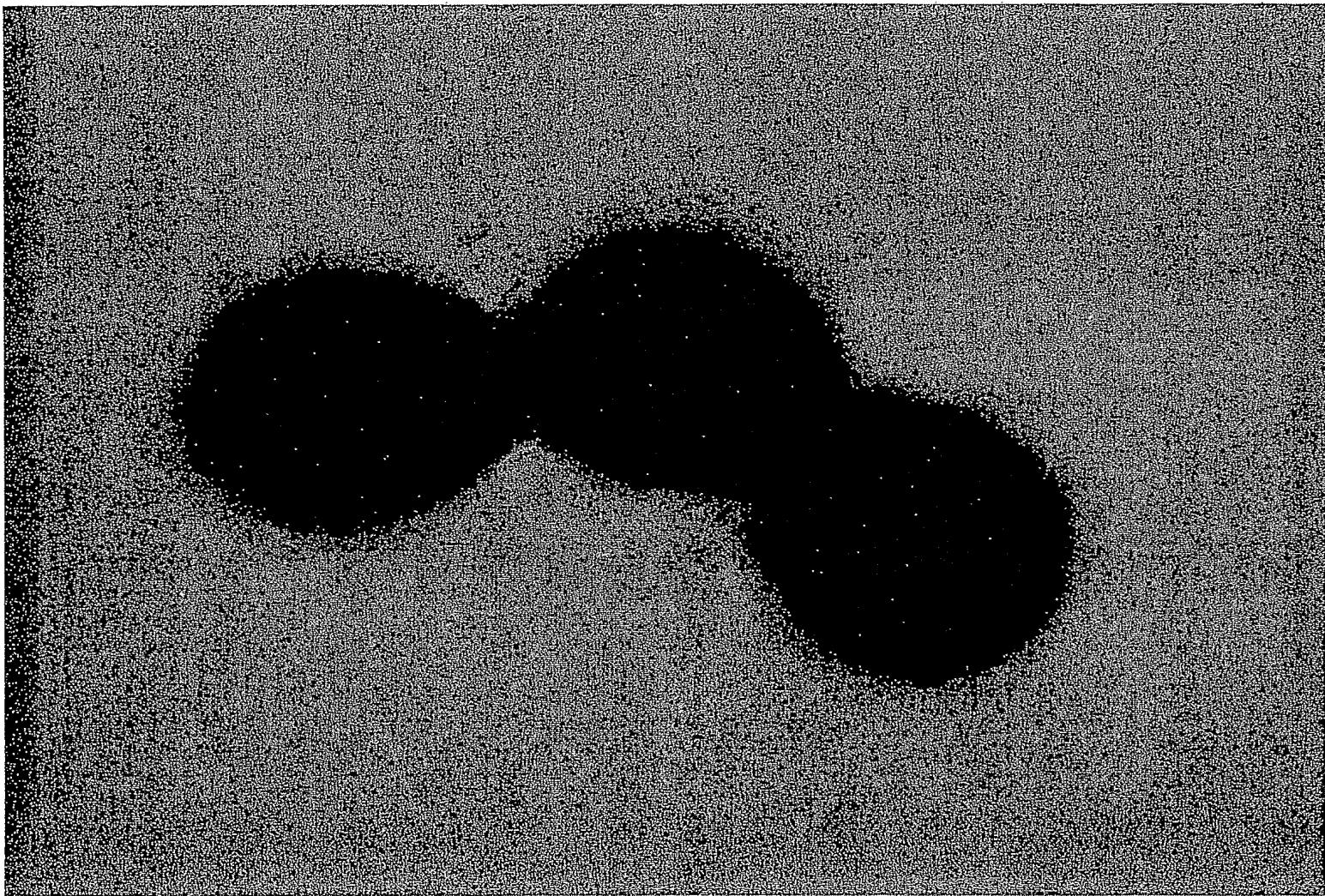


FIGURE 123

WO 2006/078318

PCT/US05/27239

PCT/US2005/027239

315/487

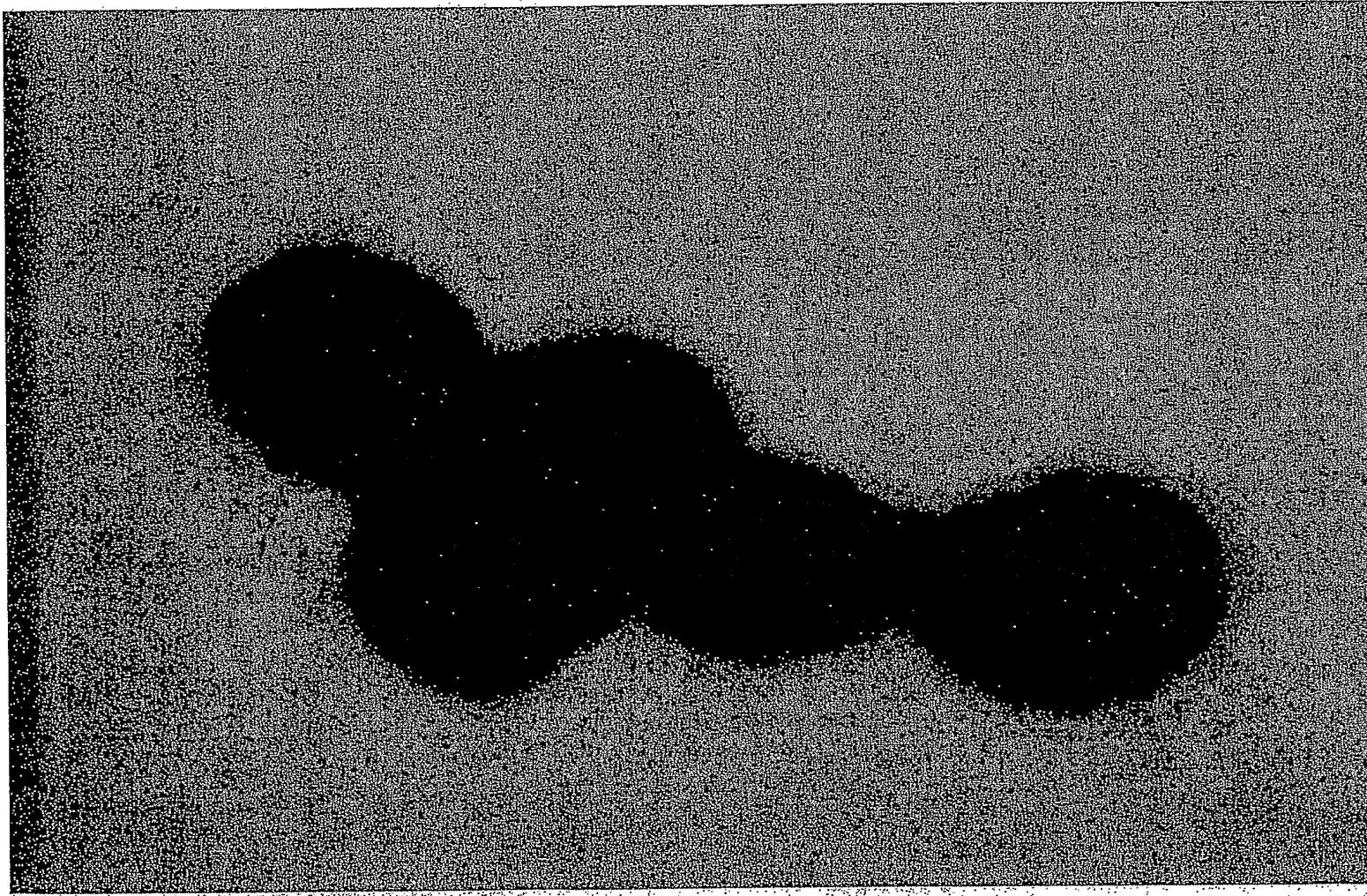


FIGURE 124

WO 2006/078318

316/487

PCT/US05/27239

PCT/US2005/027239

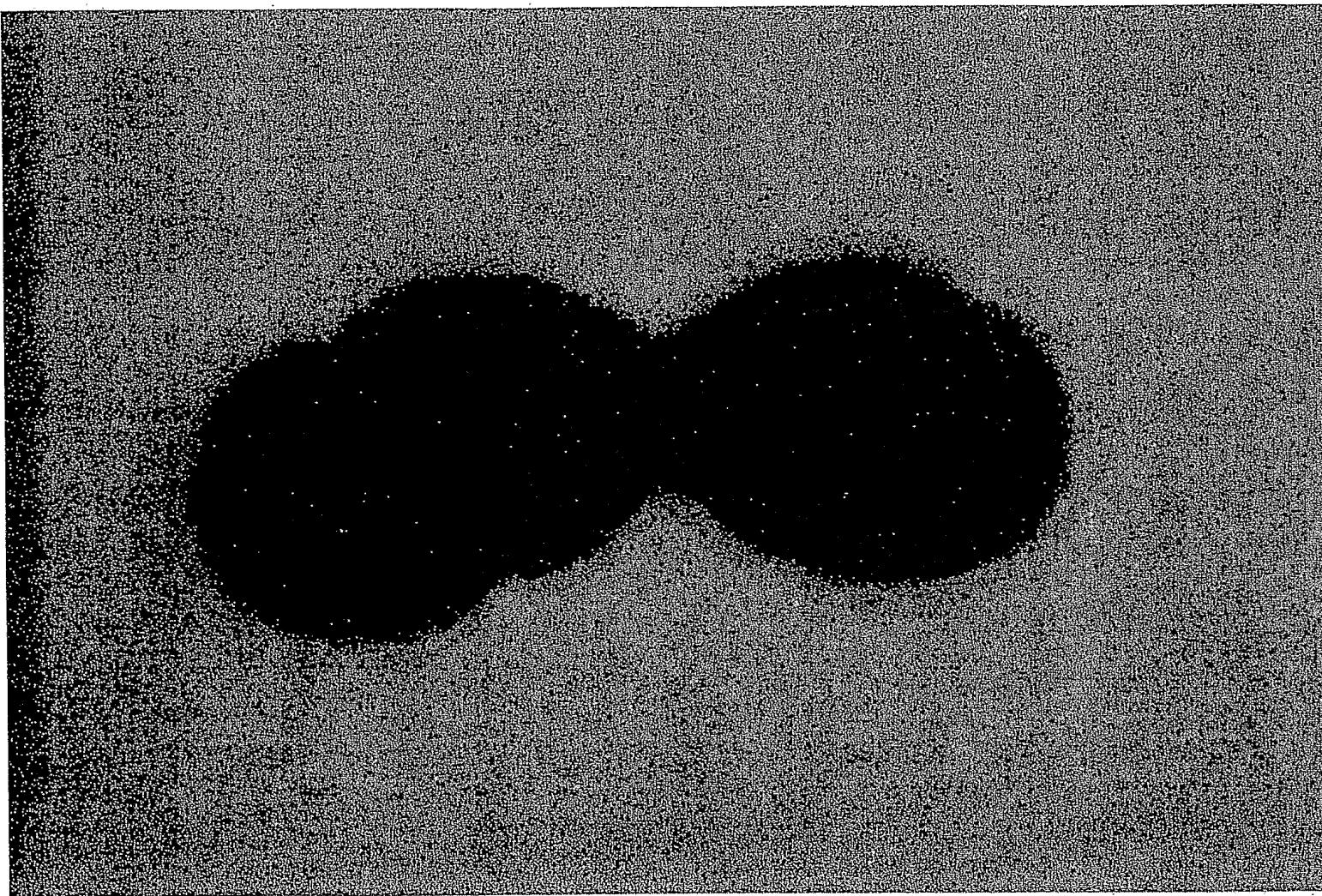


FIGURE 125

WO 2006/078318

PCT/US2005/27239

PCT/US2005/027239

317/487

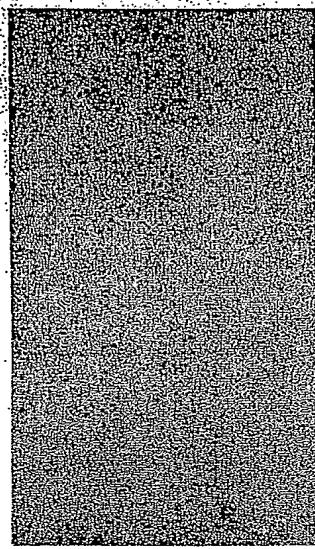


FIGURE 126

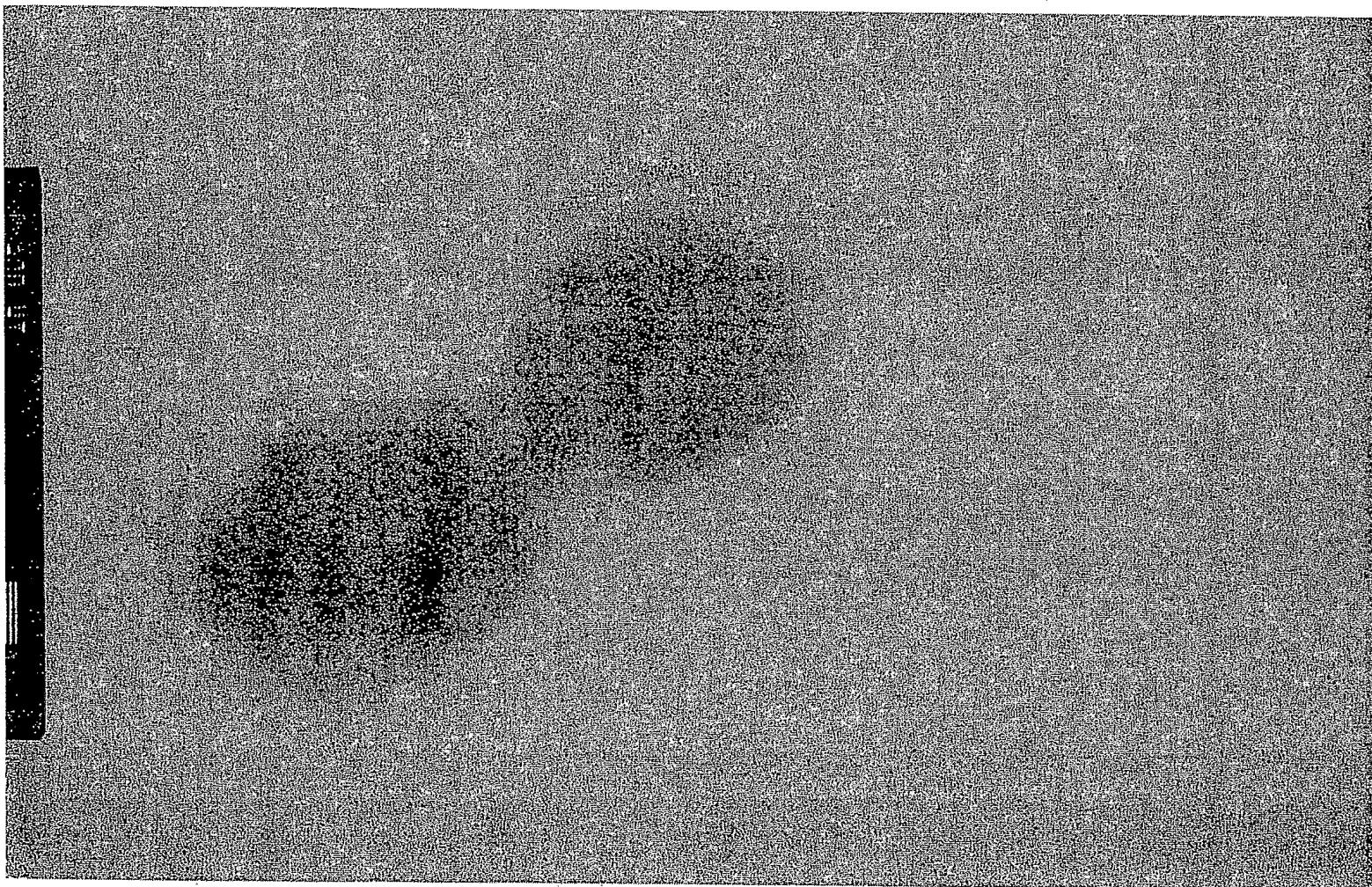
WO 2006/078318

PCT/US2005/027239

318/487

PCT/US2005/027239

Figure 127



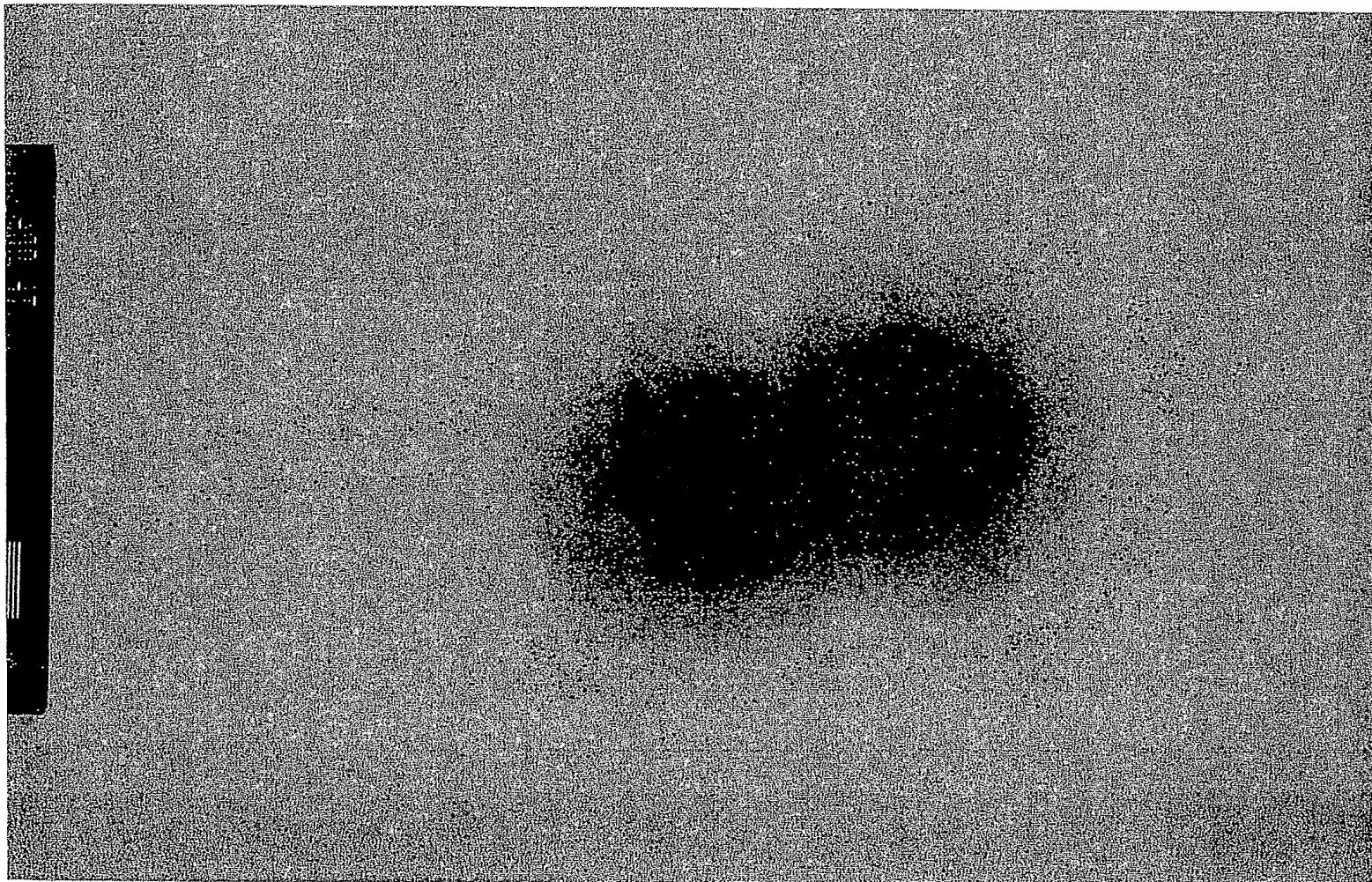
WO 2006/078318

PCT/US2005/27239

319/487

PCT/US2005/027239

Figure 128



WO 2006/078318

PCT/US2005/27239

320/487

PCT/US2005/027239

Figure 129

